(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 December 2002 (05.12.2002)

PCT

(10) International Publication Number
WO 02/097127 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/EP02/05835

(22) International Filing Date: 28 May 2002 (28.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0113266.1 31 May 2001 (31.05.2001) GB

(71) Applicant *(for all designated States except US)*: **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **OELLERS, Nadja** [DE/DE]; Steinfeldergasse 9, 50670 Köln (DE). **GEHRMANN, Mathias** [DE/DE]; Rosenstr. 68, 47918 Tönisvorst (DE). **KALLABIS, Harald** [DE/DE]; Wolfskaul 8, 51061 Köln (DE). **HALL, Roderick** [GB/GB]; 10 Acacia Close, Woodham, Surrey KT 15 3SJ (GB). **SCHULZE, Thomas** [DE/DE]; Kaiserwaldstr. 2c, 91083 Baiersdorf (DE). **KROEGEL, Claus** [DE/DE]; Falkenweg 18, 07751 Cospeda (DE).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/097127 A2

(54) Title: GENES AND PROTEINS FOR PREVENTION, PREDICTION, DIAGNOSIS, PROGNOSIS AND TREATMENT OF CHRONIC LUNG DISEASE

(57) Abstract: Genes that are differentially expressed in lung tissue of COPD patients versus lungs of normal people are disclosed. The genes provide novel methods for the prevention, prediction, diagnosis, prognosis and treatment of chronic lung disease.

GENES AND PROTEINS FOR PREVENTION, PREDICTION, DIAGNOSIS, PROGNOSIS AND TREATMENT OF CHRONIC LUNG DISEASE

TECHNICAL FIELD OF THE INVENTION

The invention relates to polynucleotides and, more particularly to genes that are differentially expressed in diseased lungs as compared to normal lung. In particular genes that are up- or downregulated in lungs of patients with Chronic Obstructive Pulmonary Disease (COPD) are disclosed. Also disclosed are methods for utilizing such genes, polynucleotides or polypeptides derived from the genes as predictive, diagnostic or prognostic markers for chronic lung disease, particularly COPD. Also disclosed are antibodies specific to the markers to be used to identify the genes or their protein products for predictive, diagnostic or prognostic purposes or to modulate their activity in order to prevent or treat lung disease. Methods of screening for modulators (activators or inhibitors) which can be used for the regulation of the genes or the polypeptides derived therefrom and preventive or therapeutic uses of these modulators are also disclosed.

BACKGROUND OF THE INVENTION

Chronic obstructive pulmonary (or airways) disease (COPD) is a condition defined physiologically as airflow obstruction that generally results from a mixture of emphysema and peripheral airway obstruction due to chronic bronchitis (1). Emphysema is characterised by destruction of alveolar walls leading to abnormal enlargement of the air spaces of the lung. Chronic bronchitis is defined clinically as the presence of chronic productive cough for three months in each of two successive years. In COPD, airflow obstruction is usually progressive and is only partially reversible. By far the most important risk factor for development of COPD is cigarette smoking, although the disease does occur in non-smokers.

Chronic inflammation of the airways is a key pathological feature of COPD (1). The inflammatory cell population comprises increased numbers of macrophages, neutrophils and $CD8^+$ lymphocytes. Inhaled irritants such as cigarette smoke activate macrophages resident in the respiratory tract as well as epithelial cells leading to release of chemokines (e.g. interleukin-8) and other chemotactic factors which act to increase the neutrophil/monocyte trafficking from the blood into the lung tissue and airways. Neutrophils and monocytes recruited into the airways can release a variety of potentially damaging mediators such as proteolytic enzymes and reactive oxygen species. Matrix degradation and emphysema, along with airway wall thickening, surfactant dysfunction and mucus hypersecretion are all potential sequelae of this inflammatory response that lead to impaired airflow and gas exchange.

Since there is a lack of the understanding of the underlying mechanisms of COPD, the therapeutic standard is not very high. COPD is the 5^{th} leading cause of death in the world and the need for drugs is extremely high. The main risk factor for development of COPD is cigarette smoking. But only a portion (up to 20%) of all smokers develop COPD, independently of their smoking habits. A genetic predisposition of a portion of smokers to develop COPD is very likely. COPD is a subgroup of the chronic lung diseases which includes also asthma and which are characterized by a chronic inflammation and/or fibrosis of the airway tissue. Many pathophysiological features are shared among these diseases. There is a need for a better understanding of the pathophysiology of chronic lung diseases on the molecular level in order to identify those genes that could serve as novel predictive, diagnostic and/ or prognostic markers and/ or as targets for preventive and/ or therapeutic drugs for chronic lung disease. Certain genes involved in the pathophysiology of chronic lung disease have been shown to be (a) upregulated in animal models of disease (108) and human patients suffering from chronic lung disease (109) or (b) downregulated in human patients suffering from chronic lung disease (110). Consequently there is a demand to identify more genes that are implicated in the disease process since novel preventive, predictive, diagnostic, prognostic and therapeutic methods can be based on them.

SUMMARY OF THE INVENTION

The present invention is based on the identification of genes that are differentially expressed in lung tissue of patients with a clear clinical evidence of COPD compared to lung tissue of individuals without any evidence of COPD. Accordingly disclosed herein are 28 genes that are differentially expressed in COPD, as well as derivatives, fragments, analogs and homologues thereof. Any of these genes is named hereinafter "COPD GENE".

It is an objective of the invention to provide methods and reagents for the prevention, prediction, diagnosis, prognosis and treatment of chronic lung disease and COPD in particular.

In one embodiment of the invention a "COPD GENE" or a gene product of a "COPD GENE" can be used to identify cells or tissue in individuals which exhibit a phenotype predisposed to disease or a diseased phenotype, thereby (a) predicting whether an individual is at risk for the development, or (b) diagnosing whether an individual is having, or (c) prognosing the progression or the outcome of the treatment of chronic lung disease and COPD in particular.

In another embodiment the invention provides methods of screening for agents which regulate the activity of a polypeptide encoded by a "COPD GENE". A test compound is contacted with a polypeptide encoded by a "COPD GENE". Binding of the test compound to the polypeptide is detected. A test compound which binds to the polypeptide is thereby identified as a potential therapeutic agent for the treatment of chronic lung disease and more particularly COPD.

In one embodiment the invention provides another method of screening for agents which regulate the activity of a polypeptide encoded by a "COPD GENE". A test compound is contacted with a polypeptide encoded by a "COPD GENE". A

biological activity mediated by the polypeptide is detected. A test compound which decreases the biological activity is thereby identified as a potential therapeutic agent for decreasing the activity of the polypeptide encoded by a "COPD GENE" in chronic lung disease and COPD in particular. A test compound which increases the biological activity is thereby identified as a potential therapeutic agent for increasing the activity of the polypeptide encoded by a "COPD GENE" " in chronic lung disease and COPD in particular.

In another embodiment the invention provides a method of screening for agents which regulate the activity of a "COPD GENE". A test compound is contacted with a "COPD GENE" polynucleotide. Binding of the test compound to the "COPD GENE" polynucleotide is detected. A test compound which binds to the polynucleotide is thereby identified as a potential therapeutic agent for regulating the activity of the "COPD GENE" in chronic lung disease and COPD in particular.

The invention thus provides "COPD GENES" which can be used to identify compounds which may act, for example, as regulators or modulators such as agonists and antagonists, partial agonist, inverse agonist, activators, co-activators and inhibitors of the polypeptide encoded by a "COPD GENE". Acordingly, the invention provides reagents and methods for regulating a "COPD GENE" polynucleotide or a polypeptide encoded by a "COPD GENE" in chronic lung disease and more particularly COPD. The regulation can be up- or downregulation. Reagents that modulate the expression, stability or amount of a "COPD GENE" polynucleotide or the activity of the polypeptide encoded by a "COPD GENE" can be a protein, a peptide, a peptidomimetic, a nucleic acid, a nucleic acid analogue (e.g. peptide nucleic acid, locked nucleic acid) or a small molecule. Methods that the expression, stability or amount of a "COPD GENE" polynucleotide or the activity of the polypeptide encoded by a "COPD GENE" can be gene replacement therapies, antisense, ribozyme and triplex nucleic acid approaches.

One embodiment of the invention provides antibodies which specifically bind to a full-length or partial "COPD GENE" for use in prevention, diagnosis, prognosis and treatment of chronic lung disease.

Yet another embodiment of the invention is the use of a reagent which specifically binds to a "COPD GENE" polynucleotide or a polypeptide encoded by a "COPD GENE" in the preparation of a medicament for the treatment of chronic lung disease and COPD in particular.

Still another embodiment of the invention is a pharmaceutical composition which includes a reagent which specifically binds to a "COPD GENE" polynucleotide or a polypeptide encoded by a "COPD GENE", and a pharmaceutically acceptable carrier.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to 28 genes that are differentially regulated in the lungs of patients with clinical evidence of COPD. "Gene or Genes" as used herein refers to the polynucleotides of SEQ ID NO. 1 to 28, as well as derivatives, fragments, analogs and homologues thereof, the polypeptides encoded thereby, the polypeptides of SEQ ID NO. 29 to 56 and the corresponding genomic transcription units which can be derived or identified with standard techniques well known in the art using the information disclosed in Tables 1 to 5. The accession numbers of the polynucleotide sequences of the SEQ IDs NO. 1 to 28 are shown in Table 3. The accession numbers of the polypeptide sequences encoded by the polynucleotide sequences are also shown in Table 3.

The invention further relates to the use of:

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and encodes a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

e) an antisense molecule targeting one of the polynucleotide sequences specified in (a) to (d);

f) a purified polypeptide encoded by a polynucleotide sequence specified in (a) to (d)

g) a purified polypeptide comprising at least one of the sequences of SEQ ID NO. 29 to 56;

h) an antibody capable of binding to one of the polynucleotide specified in (a) to (d) or a polypeptide specified in (f) and (g)

i) a reagent identified by any of the methods as specified below that modulates the amount or activity of a polynucleotide sequence specified in (a) to (d) or a polypeptide specified in (f) and (g)

for the prevention, prediction, diagnosis, prognosis and treatment of chronic lung disease.

Polynucleotides

A „COPD GENE" polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a „COPD GENE" polypeptide. Degenerate nucleotide sequences encoding human „COPD GENE" polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to the nucleotide sequences of SEQ ID NO. 1 to 28 also are „COPD GENE" polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologues, and variants of „COPD GENE" polynucleotides which encode biologically active „COPD GENE" polypeptides also are „COPD GENE" polynucleotides.

Identification of Polynucleotide Variants and Homologues

Variants and homologues of the „COPD GENE" polynucleotides described above also are „COPD GENE" polynucleotides. Typically, homologous „COPD GENE" polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known „COPD GENE" polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions: 2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50 EC once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologues of the „COPD GENE" polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression

libraries from other species, such as mice, monkeys, or yeast. Human variants of „COPD GENE" polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5°C with every 1% decrease in homology (2). Variants of human „COPD GENE" polynucleotides or „COPD GENE" polynucleotides of other species can therefore be identified by hybridizing a putative homologous „COPD GENE" polynucleotide with a polynucleotide comprising any of the nucleotide sequences of SEQ ID NO. 1 to 28 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to „COPD GENE" polynucleotides or their complements following stringent hybridization and/or wash conditions also are „COPD GENE" polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example in Ref. 3 at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20°C below the calculated T_m of the hybrid under study. The T_m of a hybrid between a „COPD GENE" polynucleotide comprising a nucleotide sequence selected from the polynucleotides of SEQ ID NO. 1 to 28 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences can be calculated, for example, using the equation of Bolton and McCarthy (4):

$$T_m = 81.5°C - 16.6(\log_{10}[Na^+]) + 0.41(\%G + C) - 0.63(\%\text{formamide}) - 600/l),$$

where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65°C, or 50% formamide, 4X SSC at 42°C, or 0.5X SSC, 0.1% SDS at 65°C. Highly stringent wash conditions include, for example, 0.2X SSC at 65°C.

Preparation of Polynucleotides

A naturally occurring „COPD GENE" polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated „COPD GENE" polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments which comprise „COPD GENE" nucleotide sequences. Isolated polynucleotides are in preparations which are free or at least 70, 80, or 90% free of other molecules.

„COPD GENE" cDNA molecules can be made with standard molecular biology techniques, using „COPD GENE" mRNA as a template. „COPD GENE" cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook et al. (3). An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesize „COPD GENE" polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a „COPD GENE" polypeptide or a biologically active variant thereof.

Extending Polynucleotides

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (5). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (6). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 2230 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (7). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker et al (8). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA (CLONTECH,

Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' nontranscribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (e.g. GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

Polypeptides

"COPD GENE" polypeptides according to the invention comprise an amino acid selected from SEQ ID NO. 29 to 56 or which are encoded by the polynucleotide sequences of SEQ ID NO. 1 to 28 or derivatives, fragments, analogs and homologues thereof. A COPD GENE" polypeptide of the invention therefore can be a portion, a full-length, or a fusion protein comprising all or a portion of a "COPD GENE" polypeptide.

The invention additionally, encompasses "COPD GENE" polypeptides which are differentially modified during or after translation, e.g. by glycosylation, acetylation, phosphorylation, amidation. derivatization by known protecting/blocking groups, proteoiytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques. including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, $NaBH_4$; acetylation, formylation. oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

Additional post-translational modifications encompassed by the invention include, for example, e.g., N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of procaryotic host cell expression. The "COPD GENE" polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

Also provided by the invention are chemically modified derivatives of the "COPD GENE" polypeptides which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent No. 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

Biologically Active Variants

„COPD GENE" polypeptide variants which are biologically active, i.e., retain an „COPD GENE" activity, also are „COPD GENE" polypeptides. Preferably, naturally or non-naturally occurring „COPD GENE" polypeptide variants have amino acid sequences which are at least about 60, 65, or 70, preferably about 75, 80, 85, 90, 92, 94, 96, or 98% identical to the amino acid sequence which is encoded by a polynucleotide comprising at least one of the polynucleotide sequences of SEQ ID NO. 1 to 28 or which comprises at least one polypeptide of SEQ ID NO. 29 to 56 or a fragment thereof. Percent identity between a putative „COPD GENE" polypeptide variant and an amino acid sequence which is encoded by a polynucleotide comprising at least one of the polynucleotide sequences of SEQ ID NO. 1 to 28 or which comprises at least one polypeptide of SEQ ID NO. 29 to 56 is determined using the Needleman/Wunsch algorithm (111) with the substitutions-matrix BLOSUM62 (112) and a gap creation penalty of 8 and a gap extension penalty of 2.

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a „COPD GENE" polypeptide can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active „COPD GENE" polypeptide can readily be determined by assaying for „COPD GENE" activity, as described for example, in the specific examples, below. Larger insertions or deletions

can also be caused by alternative splicing. Protein domains can be inserted or deleted without altering the main activity of the protein.

Fusion Proteins

Fusion proteins are useful for generating antibodies against „COPD GENE" polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins which interact with portions of a „COPD GENE" polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A „COPD GENE" polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 25, 50, 75, 100, 150, 200, 300, 400, 500, 600, 700 or 750 contiguous amino acids of an amino acid sequence encoded by a polynucleotide comprising at least one of the polynucleotide sequences of SEQ ID NO. 1 to 28 or of a biologically active variant, such as those described above. The first polypeptide segment also can comprise a full-length „COPD GENE" polypeptide.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A

fusion protein also can be engineered to contain a cleavage site located between the „COPD GENE" polypeptide-encoding sequence and the heterologous protein sequence, so that the „COPD GENE" polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from polynucleotides comprising at least one of the polynucleotide sequences of SEQ ID NO. 1 to 28 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Obtaining Polypeptides

„COPD GENE" polypeptides can be obtained, for example, by purification from human cells, by expression of „COPD GENE" polynucleotides, or by direct chemical synthesis.

Protein Purification

„COPD GENE" polypeptides can be purified from any cell which expresses the enzyme, including host cells which have been transfected with „COPD GENE" expression constructs. Lung is an especially useful source of „COPD GENE" polypeptides. A purified „COPD GENE" polypeptide is separated from other compounds which normally associate with the „COPD GENE" polypeptide in the cell, such as

certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified „COPD GENE" polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

Expression of Polynucleotides

To express a „COPD GENE" polynucleotide, the polynucleotide can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding „COPD GENE" polypeptides and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described, for example, in Sambrook et al. (3) and in Ausubel et al. (9). A variety of expression vector/host systems can be utilized to contain and express sequences encoding a „COPD GENE" polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (e.g., baculovirus), plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those nontranslated regions of the vector enhancers, promoters, 5' and 3' untranslated regions which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any

number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (e.g., heat shock, RUBISCO, and storage protein genes) or from plant viruses (e.g., viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a „COPD GENE" polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the „COPD GENE" polypeptide. For example, when a large quantity of the „COPD GENE" polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the „COPD GENE" polypeptide can be ligated into the vector in frame with sequences for the amino terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, J. Biol. Chem. 264, 55035509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or

factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel et al. (9) and Grant et al. (10).

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding „COPD GENE" polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used alone or in combination with the omega leader sequence from TMV (11). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (12,13,14). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews e.g., Hobbs or Murray (15).

An insect system also can be used to express a „COPD GENE" polypeptide. For example, in one such system Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. Sequences encoding „COPD GENE" polypeptides can be cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of „COPD GENE" polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect S. frugiperda cells or Trichoplusia larvae in which „COPD GENE" polypeptides can be expressed (16).

Mammalian Expression Systems

A number of viral-based expression systems can be used to express „COPD GENE" polypeptides in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding „COPD GENE" polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome can be used to obtain a viable virus which is capable of expressing a „COPD GENE" polypeptide in infected host cells (17). If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (e.g., liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding „COPD GENE" polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a „COPD GENE" polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used (see Ref. 18).

Host Cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed „COPD GENE“ polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Posttranslational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for Post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express „COPD GENE“ polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 12 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced „COPD GENE“ sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (19) and adenine phosphoribosyltransferase (20) genes which can be employed in tk^- or $aprt^-$ cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, dhfr confers resistance to metho-

trexate (21), npt confers resistance to the aminoglycosides, neomycin and G418 (22), and als and pat confer resistance to chlorsulfuron and phosphinotricin acetyl-transferase, respectively (Murray, 1992, supra). Additional selectable genes have been described. For example, trpB allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (23). Visible markers such as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (24).

Detecting Polynucleotide and Polypeptide Expression in Host Cells

Although the presence of marker gene expression suggests that the „COPD GENE" polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a „COPD GENE" polypeptide is inserted within a marker gene sequence, transformed cells containing sequences which encode a „COPD GENE" polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a „COPD GENE" polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the „COPD GENE" polynucleotide.

Alternatively, host cells which contain a „COPD GENE" polynucleotide and which express a „COPD GENE" polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a polynucleotide sequence encoding a „COPD GENE" polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a „COPD GENE" polypeptide.

Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a „COPD GENE" polypeptide to detect transformants which contain a „COPD GENE" polynucleotide.

A variety of protocols for detecting and measuring the expression of a „COPD GENE" polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a „COPD GENE" polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton et al. (25) and (26).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding „COPD GENE" polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a „COPD GENE" polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes in vitro by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding a „COPD GENE" polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode „COPD GENE" polypeptides can be designed to contain signal sequences which direct secretion of soluble „COPD GENE" polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound „COPD GENE" polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a „COPD GENE" polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the „COPD GENE" polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a „COPD GENE" polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described in (27) while the enterokinase cleavage site provides a means for purifying the „COPD GENE" polypeptide from the fusion protein. Vectors which contain fusion proteins are disclosed in (28).

Chemical Synthesis

Sequences encoding a „COPD GENE" polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Ref. 29, 30). Alternatively, a „COPD GENE" polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (31, 32). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of „COPD GENE" polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (e.g. Ref. 33). The composition of a synthetic „COPD GENE" polypeptide can be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; see Creighton, supra). Additionally, any portion of the amino acid sequence of the „COPD GENE" polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce „COPD GENE" polypeptide-encoding nucleotide sequences possessing non-natural occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter „COPD GENE" polypeptide-encoding sequences

for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Predictive, Diagnostic and Prognostic Methods and Reagents

The present invention provides methods and reagents for determining whether a subject is at risk for developing chronic lung disease or COPD in particular by detecting the disclosed biomarkers, i.e., the disclosed polynucleotides specified in Table 3 and/or polypeptides encoded thereby. One embodiment is a method for the prediction, diagnosis or prognosis of chronic lung disease by the detection:

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO.1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

in a biological sample comprising the following steps: hybridizing at least one polynucleotide specified in (a) to (d) to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and detecting said hybridization complex.

Another embodiment is a method as described above, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

Another embodiment is a method for the prediction, diagnosis or prognosis of chronic lung disease by the detection of:

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

e) a purified polypeptide encoded by a polynucleotide sequence specified in (a) to (d);

f) a purified polypeptide comprising at least one of the sequences of SEQ ID NO. 29 to 56;

comprising the steps of contacting a biological sample with a reagent which specifically interacts with the polynucleotide specified in (a) to (d) or the polypeptide specified in (e) and (f).

Another embodiment is a diagnostic kit for conducting any of the methods described in the embodiments above.

In clinical applications, human tissue samples can be screened for the presence and/or absence of the biomarkers („COPD GENES") identified herein. Such samples could consist of needle biopsy cores, surgical resection samples, lymph node tissue, whole blood, plasma or serum. For example, these methods include obtaining a biopsy, which is optionally fractionated by cryostat sectioning to enrich diseased cells to about 80% of the total cell population. In certain embodiments, nucleic acids extracted from these samples may be amplified using techniques well known in the art. The levels of selected markers can be detected in diseased lung and healthy lung tissue samples and compared within statistically valid reference groups. An abnormal level of the „COPD GENE" polypeptide or mRNA levels is likely to be indicative of chronic lung disease such as COPD.

In even another embodiment the diagnostic method comprises determining whether a subject has an abnormal mRNA and/or protein level of the disclosed markers, such as by Northern blot analysis, reverse transcription-polymerase chain reaction (RT-PCR), *in* situ hybridization, immunoprecipitation, Western blot analyisis, or immuno-histochemistry. According to the method, cells are obtained from a subject and the levels of the disclosed „COPD GENES", protein or mRNA level, is determined and compared.

Accordingly, in one aspect, the invention provides probes and primers that are specific to the unique polynucleotide „COPD GENES" disclosed herein. Accordingly, the polynucleotide probes comprise a nucleotide sequence at least 12 nucleotides in length, preferably at least 15 nucleotides, more preferably, 25

nucleotides, and most preferably at least 40 nucleotides, and up to all or nearly all of the coding sequence which is complementary to a portion of the coding sequence of a „COPD GENE" polynucleotide sequence comprising at least one of the sequences of SEQ ID NO. 1 to 28 or a sequence complementary thereto.

In one embodiment, the method comprises using a polynucleotide probe to determine the presence of chronic lung disease cells in a tissue from a patient. Specifically, the method comprises:

1. providing a polynucleotide probe comprising a nucleotide sequence at least 12 nucleotides in length, preferably at least 15 nucleotides, more preferably, 25 nucleotides, and most preferably at least 40 nucleotides, and up to all or nearly all of the coding sequence which is complementary to a portion of the coding sequence of a polynucleotide sequence comprising at least one of the sequences of SEQ ID NO. 1 to 28 or a sequence complementary thereto and is differentially expressed in chronic lung disease, such as COPD ;

2. obtaining a tissue sample from a patient with chronic lung disease and COPD in particular;

3. providing reference tissue sample from patients with no lung disease;

4. contacting the polynucleotide probe under stringent conditions with RNA of each of said first and reference tissue samples (e.g., in a Northem blot or in situ hybridization assay); and

5. comparing (a) the amount of hybridization of the probe with RNA of the first tissue sample, with (b) the amount of hybridization of the probe with RNA of the reference tissue samples;

wherein a statistically significant difference in the amount of hybridization with the RNA of the first tissue sample as compared to the amount of hybridization of the RNA with the reference tissue samples is indicative of chronic lung disease and COPD in particular in the first tissue sample.

In one aspect, the method comprises in situ hybridization with a probe derived from a given „COPD GENE" polynucleotide, which polynucleotide sequence comprises at least one of the polynucleotide sequences of SEQ ID NO. 1 to 28 or a sequence complementary thereto. The method comprises contacting the labeled hybridization probe with a sample of a given type of tissue from a patient potentially having chronic lung disease and COPD in particular as well as normal tissue from a person with no lung disease, and determining whether the probe labels tissue of the patient to a degree significantly different (e.g., by at least a factor of two, or at least a factor of five, or at least a factor of twenty, or at least a factor of fifty) than the degree to which normal tissue is labeled.

Another such method includes the step of providing an antibody specific for the polypeptide encoded by a polynucleotide comprising at least one of the polynucleotide sequences of SEQ ID NO. 1 to 28.

The subject invention further provides a method of determining whether a cell sample obtained from a subject possesses an abnormal amount of „COPD GENE" polypeptide which comprises (a) obtaining a cell sample from the subject, (b) quantitatively determining the amount of the „COPD GENE" polypeptide in the sample so obtained, and (c) comparing the amount of the „COPD GENE" polypeptide so determined with a known standard, so as to thereby determine whether the cell sample obtained from the subject possesses an abnormal amount of the „COPD GENE" polypeptide. Such „COPD GENE" polypeptide may be detected by immunohistochemical assays, dot-blot assays, ELISA and the like.

Immunoassays are commonly used to quantify the levels of proteins in cell samples, and many other immunoassay techniques are known in the art. The invention is not limited to a particular assay procedure, and therefore is intended to include both homogeneous and heterogeneous procedures. Exemplary immunoassays which can be conducted according to the invention include fluorescence polarization immunoassay (FPIA), fluorescence immunoassay (FIA), enzyme immunoassay (EIA), nephelometric inhibition immunoassay (NIA), enzyme linked immunosorbent assay (ELISA), and radioimmunoassay (RIA). An indicator moiety, or label group, can be attached to the subject antibodies and is selected so as to meet the needs of various uses of the method which are often dictated by the availability of assay equipment and compatible immunoassay procedures. General techniques to be used in performing the various immunoassays noted above are known to those of ordinary skill in the art.

In another embodiment, the level of the encoded product, i.e., the product encoded by a polynucleotide comprising at least one of the polynucleotides sequences of SEQ ID NO. 1 to 28 or a sequence complementary thereto, in a biological fluid (e.g., blood or urine) of a patient may be determined as a way of monitoring the level of expression of the „COPD GENE“ polynucleotide sequence in cells of that patient. Such a method would include the steps of obtaining a sample of a biological fluid from the patient, contacting the sample (or proteins from the sample) with an antibody specific for a encoded „COPD GENE“ polypeptide, and determining the amount of immune complex formation by the antibody, with the amount of immune complex formation being indicative of the level of the marker encoded product in the sample. This determination is particularly instructive when compared to the amount of immune complex formation by the same antibody in a control sample taken from a normal individual or in one or more samples previously or subsequently obtained from the same person.

Of particular importance to the subject invention is the ability to quantify the level of „COPD GENE“ polypeptide as determined by the number of cells associated with a

normal or abnormal „COPD GENE" polypeptide level. The number of cells with a particular „COPD GENE" polypeptide phenotype may then be correlated with patient prognosis. In one embodiment of the invention, the „COPD GENE" polypeptide phenotype of the lesion is determined as a percentage of cells in a biopsy which are found to have abnormally high/low levels of the „COPD GENE" polypeptide. Such expression may be detected by immunohistochemical assays, dot-blot assays, ELISA and the like.

Where tissue samples are employed, immunohistochemical staining may be used to determine the number of cells having the „COPD GENE" polypeptide phenotype. For such staining, a multiblock of tissue is taken from the biopsy or other tissue sample and subjected to proteolytic hydrolysis, employing such agents as protease K or pepsin. In certain embodiments, it may be desirable to isolate a nuclear fraction from the sample cells and detect the level of the „COPD GENE" polypeptide in the nuclear fraction.

The tissues samples are fixed by treatment with a reagent such as formalin, glutaraldehyde, methanol, or the like. The samples are then incubated with an antibody, preferably a monoclonal antibody, with binding specificity for the „COPD GENE" polypeptides. This antibody may be conjugated to a label for subsequent detection of binding. Samples are incubated for a time sufficient for formation of the immunocomplexes. Binding of the antibody is then detected by virtue of a label conjugated to this antibody. Where the antibody is unlabeled, a second labeled antibody may be employed, e.g., which is specific for the isotype of the anti-„COPD GENE" polypeptide antibody. Examples of labels which may be employed include radionuclides, fluorescers, chemiluminescers, enzymes and the like.

Where enzymes are employed, the substrate for the enzyme may be added to the samples to provide a colored or fluorescent product. Examples of suitable enzymes for use in conjugates include horseradish peroxidase, alkaline phosphatase, malate dehydrogenase and the like. Where not commercially available, such antibody-

enzyme conjugates are readily produced by techniques known to those skilled in the art.

In one embodiment, the assay is performed as a dot blot assay. The dot blot assay finds particular application where tissue samples are employed as it allows determination of the average amount of the „COPD GENE" polypeptide associated with a single cell by correlating the amount of „COPD GENE" polypeptide in a cell-free extract produced from a predetermined number of cells.

In one embodiment, the present invention also provides a method wherein polynucleotide probes are immobilized on a DNA chip in an organized array. Oligonucleotides can be bound to a solid support by a variety of processes, including lithography. For example a chip can hold up to 250,000 oligonucleotides (GeneChip, Affymetrix). These polynucleotide probes comprise a nucleotide sequence at least about 12 nucleotides in length, preferably at least about 15 nucleotides, more preferably at least about 25 nucleotides, and most preferably at least about 40 nucleotides, and up to all or nearly all of a sequence which is complementary to a portion of the coding sequence of a „COPD GENE" polynucleotide sequence selected from at least one of the polynucleotides of SEQ ID NO. 1 to 28. The present invention provides significant advantages over the available tests for various chronic lung diseases, such as COPD, because it increases the reliability of the test by providing an array of polynucleotide markers on a single chip.

The method includes obtaining tissue or cells from airways such as alveolar macrophages, e.g. by bronchoalveolar lavage (BAL). The DNA or RNA is extracted, amplified, and analyzed with a DNA chip to determine the presence of absence of the „COPD GENE" polynucleotide sequences. In one embodiment, the polynucleotide probes are spotted onto a substrate in a two-dimensional matrix or array. Samples of polynucleotides can be labeled and then hybridized to the probes. Double-stranded polynucleotides, comprising the labeled sample polynucleotides bound to probe

polynucleotides, can be detected once the unbound portion of the sample is washed away.

The probe polynucleotides can be spotted on substrates including glass, nitrocellulose, etc. The probes can be bound to the substrate by either covalent bonds or by non-specific interactions, such as hydrophobic interactions. The sample polynucleotides can be labeled using radioactive labels, fluorophores, chromophores, etc. Techniques for constructing arrays and methods of using these arrays are described in EP No. 0 799 897; PCT No. WO 97/29212; PCT No. WO 97/27317; EP No. 0 785 280; PCT No. WO 97/02357; U.S. Pat. No. 5,593,839; U.S. Pat. No. 5,578,832; EP No. 0 728 520; U.S. Pat. No. 5,599,695; EP No. 0 721 016; U.S. Pat. No. 5,556,752; PCT No. WO 95/22058; and U.S. Pat. No. 5,631,734. Further, arrays can be used to examine differential expression of genes and can be used to determine gene function. For example, arrays of the instant polynucleotide sequences can be used to determine if any of the polynucleotide sequences are differentially expressed between normal cells and cells in chronic lung disease, for example. High expression of a particular message in a cell from chronically diseased lung, which is not observed in a corresponding normal cell, can indicate a protein specific for chronic lung disease such as COPD.

In yet another embodiment, the invention contemplates using a panel of antibodies which are generated against the "COPD GENE" polypeptides of this invention, which polypeptides comprise a polynucleotide selected from SEQ ID NO. 29 to 56 or fragments thereof or are encoded by a polynucleotide which comprises a sequence selected from at least one of the polynucleotides of SEQ ID NO. 1 to 28. Such a panel of antibodies may be used as a reliable diagnostic probe for chronic lung disease and COPD in particular. The assay of the present invention comprises contacting a biopsy sample containing cells, e.g., macrophages obtained by bronchoalveolar lavage, with a panel of antibodies to one or more of the encoded products to determine the presence or absence of the "COPD GENE" polypeptides.

The diagnostic methods of the subject invention may also be employed as follow-up to treatment, e.g., quantification of the level of „COPD GENE" polypeptide may be indicative of the effectiveness of current or previously employed chronic lung disease therapies as well as the effect of these therapies upon patient prognosis.

Accordingly, the present invention makes available diagnostic assays and reagents for detecting gain and/or loss of „COPD GENE" polypeptides from a cell in order to aid in the diagnosis and phenotyping of chronic lung disorders, and COPD in particular.

The diagnostic assays described above can be adapted to be used as prognostic assays, as well. Such an application takes advantage of the sensitivity of the assays of the invention to events which take place at characteristic stages in the progression of chronic lung disease. For example, a given „COPD GENE" may be up- or down-regulated at a very early stage, perhaps before the lung is irreversibly damaged, while another „COPD GENE" may be characteristically up or down regulated only at a much later stage. Such a method could involve the steps of contacting the mRNA of a test cell with a polynucleotide probe derived from a given „COPD GENE" polynucleotide which is expressed at different characteristic levels in chronically diseased lung tissue cells and COPD tissue cells in particular at different stages of progression, and determining the approximate amount of hybridization of the probe to the mRNA of the cell, such amount being an indication of the level of expression of the gene in the cell, and thus an indication of the stage of progression of chronic lung disease and COPD; alternatively, the assay can be carried out with an antibody specific for the gene product of the given „COPD GENE" polynucleotide, contacted with the proteins of the test cell. The methods of the invention can also be used to follow the clinical course of chronic lung disease. For example, the assay of the invention can be applied to a tissue sample from a patient (e.g. macrophages from BAL); following treatment of the patient for chronic lung disease, another tissue sample is taken and the test repeated. Successful treatment will result in removal of all cells which demonstrate differential expression characteristic of the chronically

diseased lung tissue cells and COPD tissue cells in particular, or a substantial increase in expression of the gene in those cells, perhaps approaching or even surpassing normal levels.

In yet another embodiment, the invention provides methods for determining whether a subject is at risk for developing a disease, such as a predisposition to develop chronic lung disease, for example COPD, associated with an aberrant activity of at least one of the polypeptides of SEQ ID NO. 29 to 56, wherein the aberrant activity of the polypeptide is characterized by detecting the presence or absence of a genetic lesion characterized by at least one of (i) an alteration affecting the integrity of a gene encoding a „COPD GENE" polypeptides, or (ii) the mis-expression of the encoding polynucleotide. To illustrate, such genetic lesions can be detected by ascertaining the existence of at least one of (i) a deletion of one or more nucleotides from the polynucleotide sequence, (ii) an addition of one or more nucleotides to the polynucleotide sequence, (iii) a substitution of one or more nucleotides of the polynucleotide sequence, (iv) a gross chromosomal rearrangement of the polynucleotide sequence, (v) a gross alteration in the level of a messenger RNA transcript of the polynucleotide sequence, (vi) aberrant modification of the polynucleotide sequence, such as of the methylation pattern of the genomic DNA, (vii) the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene, (viii) a non-wild type level of the „COPD GENE" polypeptide, (ix) allelic loss of the gene, and/or (x) inappropriate post-translational modification of the „COPD GENE" polypeptide

The present invention provides assay techniques for detecting lesions in the encoding polynucleotide sequence. These methods include, but are not limited to, methods involving sequence analysis, Southern blot hybridization, restriction enzyme site mapping, and methods involving detection of absence of nucleotide pairing between the polynucleotide to be analyzed and a probe.

Specific diseases or disorders, e.g., genetic diseases or disorders, are associated with specific allelic variants of polymorphic regions of certain genes, which do not necessarily encode a mutated protein. Thus, the presence of a specific allelic variant of a polymorphic region of a gene in a subject can render the subject susceptible to developing a specific disease or disorder. Polymorphic regions in genes, can be identified, by determining the nucleotide sequence of genes in populations of individuals. If a polymorphic region is identified, then the link with a specific disease can be determined by studying specific populations of individuals, e.g, individuals which developed a specific disease, such as chronic lung disease like COPD. A polymorphic region can be located in any region of a gene, e.g., exons, in coding or non coding regions of exons, introns, and promoter region.

In an exemplary embodiment, there is provided a polynucleotide composition comprising a polynucleotide probe including a region of nucleotide sequence which is capable of hybridizing to a sense or antisense sequence of a gene or naturally occurring mutants thereof, or 5' or 3' flanking sequences or intronic sequences naturally associated with the subject genes or naturally occurring mutants thereof. The polynucleotide of a cell is rendered accessible for hybridization, the probe is contacted with the polynucleotide of the sample, and the hybridization of the probe to the sample polynucleotide is detected. Such techniques can be used to detect lesions or allelic variants at either the genomic or mRNA level, including deletions, substitutions, etc., as well as to determine mRNA transcript levels.

A preferred detection method is allele specific hybridization using probes overlapping the mutation or polymorphic site and having about 5, 10, 20, 25, or 30 nucleotides around the mutation or polymorphic region. In a preferred embodiment of the invention, several probes capable of hybridizing specifically to allelic variants are attached to a solid phase support, e.g., a "chip". Mutation detection analysis using these chips comprising oligonucleotides, also termed "DNA probe arrays" is described e.g., in (34). In one embodiment, a chip comprises all the allelic variants of at least one polymorphic region of a gene. The solid phase support is then contacted

with a test polynucleotide and hybridization to the specific probes is detected. Accordingly, the identity of numerous allelic variants of one or more genes can be identified in a simple hybridization experiment.

In certain embodiments, detection of the lesion comprises utilizing the probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligase chain reaction (LCR) (see, e.g. Ref. 35 and 36), the latter of which can be particularly useful for detecting point mutations in the gene (see Ref. 37). In a merely illustrative embodiment, the method includes the steps of (i) collecting a sample of cells from a patient, (ii) isolating polynucleotide (e.g., genomic, mRNA or both) from the cells of the sample, (iii) contacting the polynucleotide sample with one or more primers which specifically hybridize to a polynucleotide sequence under conditions such that hybridization and amplification of the polynucleotide (if present) occurs, and (iv) detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence (38), transcriptional amplification system (39), Q-Beta Replicase (40), or any other polynucleotide amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of polynucleotide molecules if such molecules are present in very low numbers.

In a preferred embodiment of the subject assay, mutations in, or allelic variants, of a gene from a sample cell are identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are

determined by gel electrophoresis. Moreover; the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a „COPD GENE" polypeptide. An antibody as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, $F(ab)_2$, and Fv, which are capable of binding an epitope of a „COPD GENE" polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, e.g., at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a „COPD GENE" polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody which specifically binds to the immunogen.

Typically, an antibody which specifically binds to a „COPD GENE" polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to „COPD GENE" polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a „COPD GENE" polypeptide from solution.

„COPD GENE" polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a „COPD GENE" polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (e.g., aluminum hydroxide), and surface active substances (e.g. lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially useful.

Monoclonal antibodies which specifically bind to a „COPD GENE" polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B cell hybridoma technique, and the EBV hybridoma technique (41, 42, 43, 44). In addition, techniques developed for the production of chimeric antibodies, the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (45, 46, 47). Monoclonal and other antibodies also can be humanized to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in treatment or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grafting of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies which specifically bind to a „COPD GENE" polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. Patent 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies which specifically bind to „COPD GENE“ polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobin libraries (48).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (49). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in (50). Construction of bivalent, bispecific single-chain antibodies is taught in (51). A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (52, 53).

Antibodies which specifically bind to „COPD GENE“ polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (54, 55). Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the antibodies described in WO 94/13804, also can be prepared.

Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a „COPD GENE“ polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense Oligonucleotides

Antisense oligonucleotides are nucleotide sequences which are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 6 nucleotides in length, but can be at least 7, 8, 10, 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of „COPD GENE“ gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, peptide nucleic acids (PNAs; described in US 5,714,331), locked nucleic acids (LNAs; described in WO 99/14226), or a combination of them. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters (see Ref. 56, 57, 58). Modifications of „COPD GENE“ polynucleotide expression can be obtained by designing antisense oligonucleotides which will form duplexes to the control, 5', or regulatory regions of the „COPD GENE“. Oligonucleotides derived from the transcription initiation site, e.g., between positions 10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (e.g., Ref. 59). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between an antisense oligonucleotide and the complementary sequence of a „COPD GENE" polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a „COPD GENE" polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent „COPD GENE" nucleotides, can provide sufficient targeting specificity for „COPD GENE" mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular „COPD GENE" polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a „COPD GENE" polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5' substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art (see, e.g., 60, 61, 62).

Ribozymes

Ribozymes are RNA molecules with catalytic activity (see, e.g., Ref. 63, 64, 65, 66). Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (e.g., Ref. 67). The mechanism of ribozyme action involves

sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammer-head motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The transcribed sequence of a „COPD GENE“ can be used to generate ribozymes which will specifically bind to mRNA transcribed from a „COPD GENE“ genomic locus. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (see Ref. 68). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Ref. 69).

Specific ribozyme cleavage sites within a „COPD GENE“ RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate „COPD GENE“ RNA targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing

DNA construct into cells in which it is desired to decrease „COPD GENE" expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in (67), ribozymes can be engineered so that ribozyme expression will occur in response to factors which induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

Screening Methods

The invention provides assays for screening test compounds which bind to or modulate the activity of a „COPD GENE" polypeptide or a „COPD GENE" polynucleotide. A test compound preferably binds to a „COPD GENE" polypeptide or polynucleotide. More preferably, a test compound decreases „COPD GENE" activity by at least about 10, preferably about 50, more preferably 90, or 100% relative to the absence of the test compound; or a test compound increases „COPD GENE" activity by at least 0.3, preferably by 0.5, more preferably by 1, 2 or 5 or more fold relative to the absence of the test compound.

Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or

synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the one-bead one-compound library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds (see Ref.70).

Methods for the synthesis of molecular libraries are well known in the art (see, for example, Ref. 71, 72, 73, 74, 75, 76, 77). Libraries of compounds can be presented in solution (see, e.g., Ref. 78), or on beads (79), chips (80), bacteria or spores (81), plasmids (82), or phage (83, 84, 85, 86 and 81).

High Throughput Screening

Test compounds can be screened for the ability to bind to „COPD GENE" polypeptides or polynucleotides or to affect „COPD GENE" activity or „COPD GENE" gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well, 384-well or 1536-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 5 to 500 µl. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the microwell formats.

Alternatively, Free format assays, or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by (87). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds

are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa., Nov. 710, 1995. Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described in (88). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in (89). In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

For binding assays, the test compound is preferably a small molecule which binds to and occupies, for example, the ATP/GTP binding site of the enzyme or the active site of a „COPD GENE“ polypeptide, such that normal biological activity is prevented.

Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or a „COPD GENE" polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound which is bound to a „COPD GENE" polypeptide can then be accomplished, for example, by direct counting of radioemmission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a „COPD GENE" polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a „COPD GENE" polypeptide. A microphysiometer (e.g., CytosensorJ) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a „COPD GENE" polypeptide (90).

Determining the ability of a test compound to bind to a „COPD GENE" polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) (91 and 92). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a „COPD GENE" polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, e.g., Ref. 93, 94, 95, 96, 97 and 98), to identify other proteins which bind to or interact with the „COPD GENE" polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide encoding a „COPD GENE“ polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (e.g., GAL4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact in vivo to form an protein- dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein which interacts with the „COPD GENE“ polypeptide.

It may be desirable to immobilize either a „COPD GENE“ polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either a „COPD GENE“ polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach a „COPD GENE“ polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a „COPD GENE“ polypeptide (or polynucleotide) can

be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, a „COPD GENE" polypeptide is a fusion protein comprising a domain that allows the „COPD GENE" polypeptide to be bound to a solid support. For example, glutathione S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the nonadsorbed „COPD GENE" polypeptide; the mixture is then incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a „COPD GENE" polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated „COPD GENE" polypeptides (or polynucleotides) or test compounds can be prepared from biotin NHS (N-hydroxysuccinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a „COPD GENE" polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the ATP/GTP binding site or the active site of the „COPD GENE" polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using anti-

bodies which specifically bind to a „COPD GENE" polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of a „COPD GENE" polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a „COPD GENE" polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a „COPD GENE" polypeptide or polynucleotide can be used in a cell-based assay system. A „COPD GENE" polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the test compound to a „COPD GENE" polypeptide or polynucleotide is determined as described above.

Enzyme Assays

As given in Table 3 some of the "COPD GENE" polypeptides have enzymatic activity. Test compounds can be tested for the ability to increase or decrease the activity of a „COPD GENE" polypeptide that has enzymatic activity. Enzymatic activity of a „COPD GENE" polypeptide can be measured, for these polypeptides, as known in the art. Enzyme assays can be carried out after contacting either a purified „COPD GENE" polypeptide, a cell membrane preparation, or an intact cell with a test compound.

Modulation of Gene Expression

In another embodiment, test compounds which increase or decrease „COPD GENE" gene expression are identified. A „COPD GENE" polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the „COPD GENE" polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this

comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide expression. Alternatively, when expression of the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

The level of „COPD GENE" mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a „COPD GENE" polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined in vivo, in a cell culture, or in an in vitro translation system by detecting incorporation of labeled amino acids into a „COPD GENE" polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell which expresses a „COPD GENE" polynucleotide can be used in a cell-based assay system. A „COPD GENE" polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions which can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a „COPD GENE" polypeptide, „COPD GENE" polynucleotide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a „COPD GENE" polypeptide, or mimetics, agonists, antagonists,

partial agonists, inverse agonists, activators, co-activators or inhibitors of a „COPD GENE“ polypeptide activity. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, inhalation, intravenous, intramuscular, intraarterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxy-propylmethylcellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 150 mM histidine, 0.1%2% sucrose, and 27% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

Therapeutic Indications and Methods

Therapies for treatment of COPD primarily relied upon medications to open the airways and decrease inflammation, oxygen supplementation and pulmonary rehabilitation. The advent of genomics-driven molecular target identification has opened up the possibility of identifying new lung disease-specific targets for therapeutic intervention that will provide safer, more effective treatments for chronic lung disease patients. Thus, newly discovered COPD-associated genes and their products can be tested for their role(s) in chronic lung disease and used as tools to discover and develop innovative therapies.

Genes playing important roles in any of the physiological processes outlined above can be characterized as chronic lung disease targets. Genes or gene fragments

identified through genomics can readily be expressed in one or more heterologous expression systems to produce functional recombinant proteins. These proteins are characterized in vitro for their biochemical properties and then used as tools in high-throughput molecular screening programs to identify chemical modulators of their biochemical activities. Modulators of target protein activity can be identified in this manner and subsequently tested in cellular and *in vivo* disease models for therapeutic activity. Optimization of lead compounds with iterative testing in biological models and detailed pharmacokinetic and toxicological analyses form the basis for drug development and subsequent testing in humans.

„COPD GENES" can be utilized as therapeutic targets for chronic lung disease and COPD in particular. Inhibition of the activity of the upregulated „COPD GENE" polynucleotides or polypeptides leads to a decreased activity of the „COPD GENE" polynucleotides or polypeptides, overexpressed in COPD. Activation of the downregulated „COPD GENE" polynucleotides or polypeptides leads to an enhanced activity of the „COPD GENE" polynucleotides or polypeptides, expressed at a lower level in COPD patients.

This invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a human „COPD GENE" polypeptide binding molecule) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above described screening assays for treatments as described herein.

A reagent which affects human „COPD GENE" activity can be administered to a human cell, either in vitro or in vivo, to reduce or increase human „COPD GENE" activity. The reagent preferably binds to an expression product of a human „COPD GENE" gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells which have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 µg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 µg of DNA per 16 nmole of liposome delivered to about 10^6 cells, and even more preferably about 2.0 µg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid

composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods which are standard in the art (see, for example, Ref. 99). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues in vivo using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, in (Ref. 100, 101, 102, 103, 104, 105).

Determination of a Therapeutically Effective Dose

The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases human „COPD GENE“ activity relative to the human „COPD GENE“ activity which occurs in the absence of the therapeutically effective dose.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

Therapeutic efficacy and toxicity, e.g., ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD_{50}/ED_{50}.

Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors which can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to treatment. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either ex vivo or in vivo using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, a gene gun, and DEAE- or calcium phosphate-mediated transfection.

Effective in vivo dosages of an antibody are in the range of about 5 µg to about 50 µg/kg, about 50 µg to about 5 mg/kg, about 100 µg to about 500 µg/kg of patient body weight, and about 200 to about 250 µg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective in vivo dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg, about 1 µg to about 2 mg, about 5 µg to about 500 µg, and about 20 µg to about 100 µg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides which express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a „COPD GENE“ polynucleotide or the activity of a "COPD GENE" polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a „COPD GENE“ polynucleotide or the activity of a „COPD GENE“ polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to „COPD GENE“-specific mRNA, quantitative RT-PCR, immunologic detection of a „COPD GENE“ polypeptide, or measurement of the „COPD GENE“ polypeptide activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination treatment can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such treatment, including, livestock, for example, cows, pigs, sheep, goats, horses and domestic animals, for example birds, dogs, cats, rabbits and other animals such as monkeys, and most preferably, humans.

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLES

EXAMPLE 1

Material and methods

1.a) Probe selection for the differential gene expression analysis

Three patients with clear clinical evidence of COPD and a cigarette consumption equivalent of less than 20 pack years compared to five patients with a cigarette consumption equivalent of more than 40 pack years and no evidence of COPD provide the material for the study. These two groups of smokers represent the end-points of the COPD disease spectrum.

The clinical study is based on a patient-stratification scheme which is exactly followed.

In detail, the two groups of patients are classified as follows:

Group 1: smokers without COPD

- cigarette smoking equivalent > 40 pack years (1py= 1 pack of cigarettes per day for one year)
- sex: male
- no airway obstruction (FEV1> 70%)
- normal diffusing capacity (> 85% of normal)
- thorax computed tomography (CT): no signs of emphysema

Group 2: smokers with severe COPD plus emphysema

- cigarette smoking equivalent < 20 pack years

- sex: male
- airway obstruction (FEV1< 70%) and no reversibility (< 12%)
- reduced diffusion capacity (< 60%)
- thorax CT: clear emphysema

The following groups are excluded from the study:

- patients with alpha1 anti-trypsin deficiency
- patients with serious diseases such as diabetes, cardiovascular diseases, liver or kidney diseases
- patients who have been treated with systemic or oral steroids, antibiotics or NSAIDs in the last two months.

The two different populations of smokers underwent bronchoalveolar lavages (BALs) in two lung segments of the same side. The BAL fluid was characterized and the presence of > 90% macrophages and at least $1.5x10^7$ cells were criteria for including the samples into the study. Macrophages were isolated.

1.b) Differential DNA expression profiling

Expression profiling was carried out using the Affymetrix Array Technology. Total RNA was prepared from these macrophages according to a standard protocol as described in: Expression Analysis Technical Manual, Affymetrix Inc., CA, USA, cDNA was generated, labeled and hybridized to the Affymetrix Hg-U95A DNA array containing oligonucleotides of 12626 human annotated genes (according to the manufacturer's instructions). The data have been analyzed by bioinformatics tools, using the Affymetrix software and additional filter criteria. The bioinformatics analysis is described in detail below.

28 Genes were identified to be at least 1.5 fold, differentially expressed in patients with COPD in comparison to patients without COPD. 18 genes which are up-regulated (Table 1) and 10 genes which are down-regulated (Table 2) in COPD patients have been identified.

To confirm the results obtained by the array analysis with a second independent experimental approach, these 28 genes were analyzed by real-time quantitative PCR (TaqMan), using the PRISM 7700 Sequence Detection System of PE Applied Biosystems (Perkin Elmer, Foster City, CA, USA. Here a fluorogenic probe, consisting of an oligonucleotide labeled with both a fluorescent reporter dye and a quencher dye, is included in a typical PCR. Amplification of the probe-specific product causes cleavage of the probe, generating an increase in reporter fluorescence. To standardize the amount of sample RNA, GAPDH was selected as a reference, since it was not differentially regulated in IC versus IVA patients. Primers and probes were selected using the Primer Express software and localized mostly in the 3' region of the coding sequence or in the 3' untranslated region (see Tables 4 and 5 for primer- and probe- sequences). All primer pairs were checked for specificity by conventional PCR reactions. TaqMan validation experiments were performed showing that the efficiencies of the target and the control amplifications are approximately equal which is a prerequisite for the relative quantitation of gene expression by the comparative C_T method.

To further verify the disease relevance of a selection of the upregulated genes identified in Table 1, real-time PCR (Taqman) using the PRISM 7700 Sequence Detection System was carried out as described above but with β_2-microglobulin instead of GAPDH for standardization using mRNA prepared from BAL macrophages from a second independent donor set (COPD patients, and smoking matched controls) -see Table 6 for primers. Up regulated genes were confirmed for the second donor set (Table 6) in 3 to 6 donors and by greater than 2-fold increase in expression.

1.c) Data analysis

According to Affymetrix' measurement technique a single gene expression measurement on one chip yields the *average difference* value and the *absolute call.* The *average difference* is a real number supposed to represent the expression value of that gene. The *absolute call* can take the values 'A' (absent), 'M' (marginal), or 'P' (present) and denotes the quality of a single hybridization. We used both the quantitative information given by the *average difference* and the qualitative information given by the *absolute call* to identify the genes which are differentially expressed in the COPD versus the normal population.

We calculated the differential expression E in the COPD versus the normal population as follows. Given D *average difference* values c_1, c_2, ..., c_D in the COPD population and N *average difference* values m_1, m_2, ..., m_N in the population of normals, we compute

$$E \equiv \exp\left(\frac{1}{D}\sum_{i=1}^{D}\ln(c_i) - \frac{1}{N}\sum_{i=1}^{N}\ln(m_i)\right).$$

If $c_i<L$ or $d_j<L$ for one or more values of i and j, we set these particular values equal to the positive number L. L represents the *average difference* level below which the measurements are not meaningful. Typically we set L=50.

Note that the particular computation of E allows for a correct comparison to TaqMan results as obtained by the '$\Delta\Delta C_t$ – method'.

A gene is called *up-regulated* in COPD versus normal if $E \geq 1.5$ and if the number of *absolute calls* equal to 'P' in the COPD population is greater than D/2.

A gene is called *down-regulated* in COPD versus normal if $E \leq 1.5$ and if the number of *absolute calls* equal to 'P' in the normal population is greater than N/2.

Based on the *average difference* values ci and dj we also calculated p-values by means of Wilcoxon's rank test. In this way we were able to test for the significance of the expression value difference of the genes in the COPD versus the normal population.

The final list of differentially regulated genes consists of all *up-regulated* and all *down-regulated* genes in the COPD versus the normal population. Those genes on this list which are interesting for a pharmaceutical application were finally validated by TaqMan.

2. Physiological and biochemical significance of the results

Tables 1 to 3 show the summary of the genes, including the differential expression values of array analysis and TaqMan analysis, which show an excellent correlation of gene expression levels.

All 28 sequences were previously described but have not been previously recognized as being differentially expressed in COPD patients versus non-COPD patients.

Of these 28 genes, three genes (SEQ ID NO: 7, 8, 21) belong to the family of chemokines, which appear to be involved in a variety of pro-inflammatory diseases and this makes them and their receptors very attractive therapeutic

targets. Inflammation of the lung is a key feature of COPD. In COPD, interleukin-8 has already been identified as a major up-regulated chemokine. Three G-protein coupled receptors (SEQ ID NO: 11, 12, 16) have been shown to be differentially regulated. GPCRs are involved in a variety of signal transduction pathways and have proven to be among the most successful drug targets for many indications. They have been shown to be expressed in lung tissue but were not implicated in COPD. Therefore novel GPCRs will be good targets for therapeutic intervention of respiratory diseases.

Two proteases (SEQ ID NO: 10, 24) have been recognized as differentially expressed. Different classes of proteases have been identified that have the potential to contribute to lung matrix destruction. A selective inhibitor of PDE4 isoforms has demonstrated already therapeutic utility in clinical trials of subjects with COPD (106).

Transcription factors (SEQ ID NO: 1, 3, 4, 5, 14), two signaling molecules (SEQ ID NO: 17, 20) as well as the two phosphatases (SEQ ID NO: 6, 25) were identified, all of them directly involved in signal transduction processes. Additionally two transporter (SEQ ID NO: 15, 26) and a transferase (SEQ ID NO: 18) have been identified.

EXAMPLE 2

Identification of test compounds that bind to „COPD GENE" polypeptides

A purified „COPD GENE" polypeptide comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well microtiter plates is contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. The test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a „COPD GENE" polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound which increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound is not incubated is identified as a compound which binds to a „COPD GENE" polypeptide.

EXAMPLE 3

Identification of a test compound which decreases „COPD GENE" gene expression

A test compound is administered to a culture of human cells transfected with a „COPD GENE" expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells which have not been transfected is incubated for the same time without the test compound to provide a negative control.

RNA is isolated from the two cultures as described in (107). Northern blots are prepared using 20 to 30 µg total RNA and hybridized with a ^{32}P-labeled „COPD GENE" -specific probe at 65°C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected from the complement of SEQ ID NO. 19. A test compound which decreases the „COPD GENE" -specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of „COPD GENE" gene expression.

EXAMPLE 4

Treatment of COPD in an animal model

Guinea pigs are exposed on a single occasion to tobacco smoke for 50 minutes. Animals are sacrificed between 10 minutes and 24 hour following the end of the

exposure and their lungs placed in RNAlater™. The lung tissue is homogenized and total RNA is extracted using a Qiagens RNeasy™ Maxi kit. Molecular Probes RiboGreen™ RNA quantitation method is used to quantify the amount of RNA in each sample. Total RNA is reverse transcribed and the resultant cDNA was used in a real-time polymerase chain reaction (PCR). The cDNA is added to a solution containing the sense and anti-sense primers and the 6-carboxy-tetramethyl-rhodamine labeled probe of the "COPD GENE". Cyclophilin is used as the housekeeping gene. The expression of the "COPD GENE" is measured using the TaqMan real-time PCR system that generates an amplification curve for each sample. From this curve a threshold cycle value is calculated: the fractional cycle number at which the amount of amplified target reaches a fixed threshold. A sample containing many copies of the "COPD GENE" will reach this threshold earlier than a sample containing fewer copies. The threshold is set at 0.2 and the threshold cycle C_T is calculated from the amplification curve. The C_T value for the "COPD GENE" is normalized using the C_T value for the housekeeping gene.

Expression of the "COPD GENE" is increased by at least 1,5-fold between 10 minutes and 3 hours post tobacco smoke exposure compared to air exposed control animals.

Test compounds are evaluated as follows. Animals are pre-treated with a test compound between 5 minutes and 1 hour prior to the tobacco smoke exposure and they are then sacrificed up to 3 hours after the tobacco smoke exposure has been completed. Control animals are pre-treated with the vehicle of the test compound via the route of administration chosen for the test compound. A test compound that reduces the tobacco smoke induced upregulation of the "COPD GENE" relative to the expression seen in vehicle treated tobacco smoke exposed animals is identified as an inhibitor of "COPD GENE" expression.

References

1. Barnes, P.J. Mechanisms in COPD. Differences from asthma. Chest 2000, 117:10S14S.
2. Bonner et al., J. Mol. Biol. 81, 123 (1973).
3. Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989.
4. Bolton and McCarthy, Proc. Natl. Acad. Sci. U.S.A. 48, 1390 (1962).
5. Sarkar, PCR Methods Applic. 2, 318322, 1993.
6. Triglia et al., Nucleic Acids Res. 16, 8186, 1988.
7. Lagerstrom et al., PCR Methods Applic. 1, 111119, 1991.
8. Parker et al., Nucleic Acids Res. 19, 30553060, 1991.
9. Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.
10. Grant et al., Methods Enzymol. 153, 516544, 1987.
11. Takamatsu, EMBO J. 6, 307311, 1987.
12. Coruzzi et al., EMBO J. 3, 16711680, 1984.
13. Broglie et al., Science 224, 838843, 1984.
14. Winter et al., Results Probl. Cell Differ. 17, 85105, 1991.
15. MCGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191196, 1992.
16. Engelhard et al., Proc. Nat. Acad. Sci. 91, 32243227, 1994.
17. Logan & Shenk, Proc. Natl. Acad. Sci. 81, 36553659, 1984.
18. Scharf et al., Results Probl. Cell Differ. 20, 125162, 1994.
19. Wigler et al., Cell 11, 22332, 1977.
20. Lowy et al., Cell 22, 81723, 1980.
21. Wigler et al., Proc. Natl. Acad. Sci. 77, 356770, 1980.
22. Colbere-Garapin et al., J. Mol. Biol. 150, 114, 1981.
23. Hartman & Mulligan, Proc. Natl. Acad. Sci. 85, 804751, 1988.
24. Rhodes et al., Methods Mol. Biol. 55, 121131, 1995.

25. Hampton et al., SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990.
26. Maddox et al., J. Exp. Med. 158, 12111216, 1983.
27. Porath et al., Prot. Exp. Purif. 3, 263281, 1992.
28. Kroll et al., DNA Cell Biol. 12, 441453, 1993.
29. Caruthers et al., Nucl. Acids Res. Symp. Ser. 215223, 1980.
30. Horn et al. Nucl. Acids Res. Symp. Ser. 225232, 1980.
31. Merrifield, J. Am. Chem. Soc. 85, 21492154, 1963.
32. Roberge et al., Science 269, 202204, 1995.
33. Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH and Co., New York, N.Y., 1983.
34. Cronin et al. (1996) Human Mutation 7:244.
35. Landegran *et al.* (1988) *Science* 241:1077-1080.
36. Nakazawa *et al.* (1994) PNAS 91:360-364.
37. Abravaya et al. (1995) *Nuc Acid Res* 23:675-682.
38. Guatelli, J.C. *et* al., 1990, Proc. Natl. Acad. Sci. USA 87:1874-1878.
39. Kwoh, D.Y. et al., 1989, Proc. Natl. Acad. Sci. USA 86:1173 1177.
40. Lizardi, P.M. *et* al., 1988, Bio/Technology 6:1197.
41. Kohler et al., Nature 256, 495497, 1985.
42. Kozbor et al., J. Immunol. Methods 81, 3142, 1985.
43. Cote et al., Proc. Natl. Acad. Sci. 80, 20262030, 1983.
44. Cole et al., Mol. Cell Biol. 62, 109120, 1984.
45. Morrison et al., Proc. Natl. Acad. Sci. 81, 68516855, 1984.
46. Neuberger et al., Nature 312, 604608, 1984.
47. Takeda et al., Nature 314, 452454, 1985.
48. Burton, Proc. Natl. Acad. Sci. 88, 1112023, 1991.
49. Thirion et al., 1996, Eur. J. Cancer Prev. 5, 507-11.
50. Coloma & Morrison, 1997, Nat. Biotechnol. 15, 159-63.
51. Mallender & Voss, 1994, J. Biol. Chem. 269, 199-206.
52. Verhaar et al., 1995, Int. J. Cancer 61, 497-501.
53. Nicholls et al., 1993, J. Immunol. Meth. 165, 81-91.

54. Orlandi et al., Proc. Natl. Acad. Sci. 86, 3833 3837, 1989.
55. Winter et al., Nature 349, 293299, 1991.
56. Brown, Meth. Mol. Biol. 20, 18, 1994.
57. Sonveaux, Meth. Mol. Biol. 26, 1-72, 1994.
58. Uhlmann et al., Chem. Rev. 90, 543583, 1990.
59. Gee et al., in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Publishing Co., Mt. Kisco, N.Y., 1994.
60. Agrawal et al., Trends Biotechnol. 10, 152158, 1992.
61. Uhlmann et al., Chem. Rev. 90, 543584, 1990.
62. Uhlmann et al., Tetrahedron. Lett. 215, 35393542, 1987.
63. Cech, Science 236, 15321539; 1987.
64. Cech, Ann. Rev. Biochem. 59, 543568; 1990.
65. Cech, Curr. Opin. Struct. Biol. 2, 605609; 1992.
66. Couture & Stinchcomb, Trends Genet. 12, 510515, 1996.
67. Haseloff et al., U.S. Patent 5,641,673.
68. Haseloff et al. Nature 334, 585591, 1988.
69. Gerlach et al., EP 321,201.
70. Lam, Anticancer Drug Des. 12, 145, 1997.
71. DeWitt et al., Proc. Natl. Acad. Sci. U.S.A. 90, 6909, 1993.
72. Erb et al. Proc. Natl. Acad. Sci. U.S.A. 91, 11422, 1994.
73. Zuckermann et al., J. Med. Chem. 37, 2678, 1994.
74. Cho et al., Science 261, 1303, 1993.
75. Carell et al., Angew. Chem. Int. Ed. Engl. 33, 2059, 1994.
76. Carell et al., Angew. Chem. Int. Ed. Engl. 33, 2061
77. Gallop et al., J. Med. Chem. 37, 1233, 1994.
78. Houghten, BioTechniques 13, 412421, 1992.
79. Lam, Nature 354, 8284, 1991.
80. Fodor, Nature 364, 555556, 1993.
81. Ladner, U.S. Patent 5,223,409.
82. Cull et al., Proc. Natl. Acad. Sci. U.S.A. 89, 18651869, 1992.
83. Scott & Smith, Science 249, 386390, 1990.

84. Devlin, Science 249, 404406, 1990.
85. Cwirla et al., Proc. Natl. Acad. Sci. 97, 63786382, 1990.
86. Felici, J. Mol. Biol. 222, 301310, 1991.
87. Jayawickreme et al., Proc. Natl. Acad. Sci. U.S.A. 19, 161418, 1994.
88. Salmon et al., Molecular Diversity 2, 5763, 1996.
89. Beutel et al., U.S. Patent 5,976,813.
90. McConnell et al., Science 257, 19061912, 1992.
91. Sjolander & Urbaniczky, Anal. Chem. 63, 23382345, 1991.
92. Szabo et al., Curr. Opin. Struct. Biol. 5, 699705, 1995.
93. U.S. Patent 5,283,317.
94. Zervos et al., Cell 72, 223232, 1993.
95. Madura et al., J. Biol. Chem. 268, 1204612054, 1993.
96. Bartel et al., BioTechniques 14, 920924, 1993.
97. Iwabuchi et al., Oncogene 8, 16931696, 1993.
98. Brent W094/10300.
99. U.S. Patent 5,705,151.
100. Findeis et al. Trends in Biotechnol. 11, 202-05, 1993.
101. Chiou et al., GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.), 1994.
102. Wu & Wu, J. Biol. Chem. 263, 621-24, 1988.
103. Wu et al., J. Biol. Chem. 269, 542-46, 1994.
104. Zenke et al., Proc. Natl. Acad. Sci. U.S.A. 87, 3655-59, 1990.
105. Wu et al., J. Biol. Chem. 266, 338-42, 1991.
106. Barnette & Underwood, Curr Opin Pulm Med. 6,164-9, 2000.
107. Chirgwin et al., Biochem. 18, 5294-99, 1979.
108. Walter et al., J. Exp. Med. 193, 339-52, 2001
109. Cowburn et al., J. Clin. Invest. 101, 834-46, 1998
110. Pilette et al., Am. J. Crit. Care Med. 163, 185-94, 2001
111. Needleman SB, Wunsch, J Mol Biol. 48, 443-53, 1970
112. Henikoff S, Henikoff JG, Proc. Natl. Acad. Sci. USA 89,10915-10919, 1992

SEQUENCE LISTING

<110> Bayer AG

<120> GENES AND PROTEINS FOR PREVENTION, PREDICTION, DIAGNOSIS, PROGNOSIS AND TREATMENT OF CHRONIC LUNG DISEASE

<130> LeA 35183 Foreign Countries

<150> GB 0113266.1

<151> 2001-05-31

<160> 182

<170> PatentIn version 3.1

<210> 1

<211> 1345

<212> DNA

<213> Homo sapiens

<400> 1

aaaacagccg gggctccagc gggagaacga taatgcaaag tgctatgttc ttggctgttc 60

aacacgactg cagacccatg gacaagagcg caggcagtgg ccacaagagc gaggagaagc 120

gagaaaagat gaaacggacc cttttaaaag attggaagac ccgtttgagc tacttcttac 180

aaaattcctc tactcctggg aagcccaaaa ccggcaaaaa aagcaaacag caagctttca 240

tcaagccttc tcctgaggaa gcacagctgt ggtcagaagc atttgacgag ctgctagcca 300

gcaaatatgg tcttgctgca ttcagggctt ttttaaagtc ggaattctgt gaagaaaata 360

ttgaattctg gctggcctgt gaagacttca aaaaaaccaa atcaccccaa aagctgtcct 420

caaaagcaag gaaaatatat actgacttca tagaaaagga agctccaaaa gagataaaca 480

tagattttca aaccaaaact ctgattgccc agaatataca agaagctaca agtggctgct 540

ttacaactgc ccagaaaagg gtatacagct tgatggagaa caactcttat cctcgtttct 600
tggagtcaga attctaccag gacttgtgta aaaagccaca aatcaccaca gagcctcatg 660
ctacatgaaa tgtaaaaggg agcccagaaa tggaggacat ttcattcttt ttcctgaggg 720
gaaggactgt gacctgccat aaagactgac cttgaattca gcctgggtgt tcaggaaaca 780
tcactcagaa ctattgattc aaagttgggt agtgaatcag gaagccagta actgactagg 840
agaagctggt atcagaacag cttccctcac tgtgtacaga acgcaagaag ggaataggtg 900
gtctgaacgt ggtgtctcac tctgaaaagc aggaatgtaa gatgatgaaa gagacaatgt 960
aatactgttg gtccaaaagc atttaaaatc aatagatctg ggattatgtg gccttaggta 1020
gctggttgta catctttccc taaatcgatc catgttacca catagtagtt ttagtttagg 1080
attcagtaac agtgaagtgt ttactatgtg caagggtatt gaagttctta tgaccacaga 1140
tcatcagtac tgttgtctca tgtaatgcta aaactgaaat ggtccgtgtt tgcattgtta 1200
aaaatgatgt gtgaaataga atgagtgcta tggtgttgaa aactgcagtg tccgttatga 1260
gtgccaaaaa tctgtcttga aggcagctac actttgaagt ggtctttgaa tacttttaat 1320
aaatttattt tgataaataa tattg 1345

<210> 2
<211> 5722
<212> DNA
<213> Homo sapiens

<400> 2
ggacgcacag gcattccccg cgcccctcca gccctcgccg ccctcgccac cgctcccggc 60
cgccgcgctc cggtacacac aggatccctg ctgggcacca acagctccac catggggctg 120
gcctggggac taggcgtcct gttcctgatg catgtgtgtg gcaccaaccg cattccagag 180
tctggcggag acaacagcgt gtttgacatc tttgaactca ccggggccgc ccgcaagggg 240
tctgggcgcc gactggtgaa gggccccgac ccttccagcc cagctttccg catcgaggat 300
gccaacctga tccccccctgt gcctgatgac aagttccaag acctggtgga tgctgtgcgg 360
gcagaaaagg gtttcctcct tctggcatcc ctgaggcaga tgaagaagac ccggggcacg 420
ctgctggccc tggagcggaa agaccactct ggccaggtct tcagcgtggt gtccaatggc 480
aaggcgggca ccctggacct cagcctgacc gtccaaggaa agcagcacgt ggtgtctgtg 540

gaagaagctc tcctggcaac cggccagtgg aagagcatca ccctgtttgt gcaggaagac 600
agggcccagc tgtacatcga ctgtgaaaag atggagaatg ctgagttgga cgtccccatc 660
caaagcgtct tcaccagaga cctggccagc atcgccagac tccgcatcgc aaaggggggc 720
gtcaatgaca atttccaggg ggtgctgcag aatgtgaggt ttgtctttgg aaccacacca 780
gaagacatcc tcaggaacaa aggctgctcc agctctacca gtgtcctcct caccccttgac 840
aacaacgtgg tgaatggttc cagccctgcc atccgcacta actacattgg ccacaagaca 900
aaggacttgc aagccatctg cggcatctcc tgtgatgagc tgtccagcat ggtcctggaa 960
ctcaggggcc tgcgcaccat tgtgaccacg ctgcaggaca gcatccgcaa agtgactgaa 1020
gagaacaaag agttggccaa tgagctgagg cggcctcccc tatgctatca caacggagtt 1080
cagtacagaa ataacgagga atggactgtt gatagctgca ctgagtgtca ctgtcagaac 1140
tcagttacca tctgcaaaaa ggtgtcctgc cccatcatgc cctgctccaa tgccacagtt 1200
cctgatggag aatgctgtcc tcgctgttgg cccagcgact ctgcggacga tggctggtct 1260
ccatggtccg agtggacctc ctgttctacg agctgtggca atggaattca gcagcgcggc 1320
cgctcctgcg atagcctcaa caaccgatgt gagggctcct cggtccagac acggacctgc 1380
cacattcagg agtgtgacaa aagatttaaa caggatggtg gctggagcca ctggtccccg 1440
tggtcatctt gttctgtgac atgtggtgat ggtgtgatca caaggatccg gctctgcaac 1500
tctcccagcc cccagatgaa tgggaaaccc tgtgaaggcg aagcgcggga gaccaaagcc 1560
tgcaagaaag acgcctgccc catcaatgga ggctggggtc cttggtcacc atgggacatc 1620
tgttctgtca cctgtggagg aggggtacag aaacgtagtc gtctctgcaa caaccccgca 1680
ccccagtttg gaggcaagga ctgcgttggt gatgtaacag aaaaccagat ctgcaacaag 1740
caggactgtc caattgatgg atgcctgtcc aatccctgct ttgccggcgt gaagtgtact 1800
agctaccctg atggcagctg gaaatgtggt gcttgtcccc ctggttacag tggaaatggc 1860
atccagtgca cagatgttga tgagtgcaaa gaagtgcctg atgcctgctt caaccacaat 1920
ggagagcacc ggtgtgagaa cacggacccc ggctacaact gcctgccctg ccccccacgc 1980
ttcaccggct cacagccctt cggccagggt gtcgaacatg ccacggccaa caaacaggtg 2040
tgcaagcccc gtaacccctg cacggatggg acccacgact gcaacaagaa cgccaagtgc 2100
aactacctgg gccactatag cgaccccatg taccgctgcg agtgcaagcc tggctacgct 2160
ggcaatggca tcatctgcgg ggaggacaca gacctggatg gctggcccaa tgagaacctg 2220
gtgtgcgtgg ccaatgcgac ttaccactgc aaaaaggata attgccccaa ccttcccaac 2280
tcagggcagg aagactatga caaggatgga attggtgatg cctgtgatga tgacgatgac 2340

aatgataaaa ttccagatga cagggacaac tgtccattcc attacaaccc agctcagtat 2400
gactatgaca gagatgatgt gggagaccgc tgtgacaact gtccctacaa ccacaaccca 2460
gatcaggcag acacagacaa caatggggaa ggagacgcct gtgctgcaga cattgatgga 2520
gacggtatcc tcaatgaacg ggacaactgc cagtacgtct acaatgtgga ccagagagac 2580
actgatatgg atggggttgg agatcagtgt gacaattgcc ccttggaaca caatccggat 2640
cagctggact ctgactcaga ccgcattgga gatacctgtg acaacaatca ggatattgat 2700
gaagatggcc accagaacaa tctggacaac tgtccctatg tgcccaatgc caaccaggct 2760
gaccatgaca aagatggcaa gggagatgcc tgtgaccacg atgatgacaa cgatggcatt 2820
cctgatgaca aggacaactg cagactcgtg cccaatcccg accagaagga ctctgacggc 2880
gatggtcgag gtgatgcctg caaagatgat tttgaccatg acagtgtgcc agacatcgat 2940
gacatctgtc ctgagaatgt tgacatcagt gagaccgatt tccgccgatt ccagatgatt 3000
cctctggacc ccaaagggac atcccaaaat gaccctaact gggttgtacg ccatcagggt 3060
aaagaactcg tccagactgt caactgtgat cctggactcg ctgtaggtta tgatgagttt 3120
aatgctgtgg acttcagtgg caccttcttc atcaacaccg aaagggacga tgactatgct 3180
ggatttgtct ttggctacca gtccagcagc cgcttttatg ttgtgatgtg gaagcaagtc 3240
acccagtcct actgggacac caaccccacg agggctcagg gatactcggg cctttctgtg 3300
aaagttgtaa actccaccac agggcctggc gagcacctgc ggaacgccct gtggcacaca 3360
ggaaacaccc ctggccaggt gcgcaccctg tggcatgacc ctcgtcacat aggctggaaa 3420
gatttcaccg cctacagatg gcgtctcagc cacaggccaa agacgggttt cattagagtg 3480
gtgatgtatg aagggaagaa aatcatggct gactcaggac ccatctatga taaaacctat 3540
gctggtggta gactagggtt gtttgtcttc tctcaagaaa tggtgttctt ctctgacctg 3600
aaatacgaat gtagagatcc ctaatcatca aattgttgat tgaaagactg atcataaacc 3660
aatgctggta ttgcaccttc tggaactatg ggcttgagaa aacccccagg atcacttctc 3720
cttggcttcc ttcttttctg tgcttgcatc agtgtggact cctagaacgt gcgacctgcc 3780
tcaagaaaat gcagttttca aaaacagact catcagcatt cagcctccaa tgaataagac 3840
atcttccaag catataaaca attgctttgg tttccttttg aaaaagcatc tacttgcttc 3900
agttgggaag gtgcccattc cactctgcct ttgtcacaga gcagggtgct attgtgaggc 3960
catctctgag cagtggactc aaaagcattt tcaggcatgt cagagaaggg aggactcact 4020
agaattagca aacaaaacca ccctgacatc ctccttcagg aacacgggga gcagaggcca 4080

aagcactaag gggagggcgc atacccgaga cgattgtatg aagaaaatat ggaggaactg 4140
ttacatgttc ggtactaagt cattttcagg ggattgaaag actattgctg gatttcatga 4200
tgctgactgg cgttagctga ttaacccatg taaataggca cttaaataga agcaggaaag 4260
ggagacaaag actggcttct ggacttcctc cctgatcccc acccttactc atcaccttgc 4320
agtggccaga attagggaat cagaatcaaa ccagtgtaag gcagtgctgg ctgccattgc 4380
ctggtcacat tgaaattggt ggcttcattc tagatgtagc ttgtgcagat gtagcaggaa 4440
aataggaaaa cctaccatct cagtgagcac cagctgcctc ccaaaggagg ggcagccgtg 4500
cttatatttt tatggttaca atggcacaaa attattatca acctaactaa aacattcctt 4560
ttctcttttt tccgtaatta ctaggtagtt ttctaattct ctcttttgga agtatgattt 4620
ttttaaagtc tttacgatgt aaaatattta ttttttactt attctggaag atctggctga 4680
aggattattc atggaacagg aagaagcgta aagactatcc atgtcatctt tgttgagagt 4740
cttcgtgact gtaagattgt aaatacagat tatttattaa ctctgttctg cctggaaatt 4800
taggcttcat acggaaagtg tttgagagca agtagttgac atttatcagc aaatctcttg 4860
caagaacagc acaaggaaaa tcagtctaat aagctgctct gccccttgtg ctcagagtgg 4920
atgttatggg attccttttt tctctgtttt atcttttcaa gtggaattag ttggttatcc 4980
atttgcaaat gttttaaatt gcaaagaaag ccatgaggtc ttcaatactg ttttacccca 5040
tcccttgtgc atatttccag ggagaaggaa agcatataca ctttttttctt tcatttttcc 5100
aaaagagaaa aaaatgacaa aaggtgaaac ttacatacaa atattacctc atttgttgtg 5160
tgactgagta aagaattttt ggatcaagcg gaaagagttt aagtgtctaa caaacttaaa 5220
gctactgtag tacctaaaaa gtcagtgttg tacatagcat aaaaactctg cagagaagta 5280
ttcccaataa ggaaatagca ttgaaatgtt aaatacaatt tctgaaagtt atgttttttt 5340
tctatcatct ggtataccat tgctttattt ttataaatta ttttctcatt gccattggaa 5400
tagaatattc agattgtgta gatatgctat ttaaataatt tatcaggaaa tactgcctgt 5460
agagttagta tttctatttt tatataatgt ttgcacactg aattgaagaa ttgttggttt 5520
tttctttttt ttgtttttttt ttttttttttt ttttttttttg cttttgacct cccatttta 5580
ctatttgcca ataccttttt ctaggaatgt gcttttttttt gtacacattt ttatccattt 5640
tacattctaa agcagtgtaa gttgtatatt actgtttctt atgtacaagg aacaacaata 5700
aatcatatgg aaatttatat tt 5722

<210> 3

<211> 6210

<212> DNA

<213> Homo sapiens

<400> 3
gcaggaacag tgctagtatt gctcgagccc gagggctgga ggttagggga tgaaggtctg 60

cttccacgct ttgcactgaa ttagggctag aattggggat gggggtaggg gcgcattcct 120

tcgggagccg aggcttaagt cctcggggtc ctgtactcga tgccgtttct cctatctctg 180

agcctcagaa ctgtcttcag tttccgtaca agggtaaaaa ggcgctctct gccccatccc 240

ccccgacctc gggaacaagg gtccgcattg aaccaggtgc gaatgttctc tctcattctg 300

cgccgttccc gcctcccctc ccccagccgc ggcccccgcc tccccccgca ctgcaccctc 360

ggtgttggct gcagcccgcg agcagttccc gtcaatccct ccccccttac acaggatgtc 420

catattagga catctgcgtc agcaggtttc cacggccttt ccctgtagcc ctgggggggag 480

ccatccccga aacccctcat cttgggggggc ccacgagacc tctgagacag gaactgcgaa 540

atgctcacga gattaggaca cgcgccaagg cgggggcagg gagctgcgag cgctggggac 600

gcagccgggc ggccgcagaa gcgcccaggc ccgcgcgcca cccctctggc gccaccgtgg 660

ttgagcccgt gacgtttaca ctcattcata aaacgcttgt tataaaagca gtggctgcgg 720

cgcctcgtac tccaaccgca tctgcagcga gcaactgaga agccaagact gagccggcgg 780

ccgcggcgca gcgaacgagc agtgaccgtg ctcctaccca gctctgcttc acagcgccca 840

cctgtctccg cccctcggcc cctcgcccgg ctttgcctaa ccgccacgat gatgttctcg 900

ggcttcaacg cagactacga ggcgtcatcc tcccgctgca gcagcgcgtc cccggccggg 960

gatagcctct cttactacca ctcacccgca gactccttct ccagcatggg ctcgcctgtc 1020

aacgcgcagg taaggctggc ttcccgtcgc cgcggggccg ggggcttggg gtcgcggagg 1080

aggagacacc gggcgggacg ctccagtaga tgagtagggg gctcccttgt gcctggaggg 1140

aggctgccgt ggccggagcg gtgccggctc gggggctcgg gacttgctct gagcgcacgc 1200

acgcttgcca tagtaagaat tggttccccc ttcgggaggc aggttcgttc tgagcaacct 1260

ctggtctgca ctccaggacg gatctctgac attagctgga gcagacgtgt cccaagcaca 1320

aactcgctaa ctagagcctg gcttcttcgg ggaggtggca gaaagcggca atccccctc 1380

ccccggcagc ctggagcacg gaggagggat gagggaggag ggtgcagcgg gcgggtgtgt 1440

aaggcagttt cattgataaa aagcgagttc attctggaga ctccggagcg gcgcctgcgt 1500
cagcgcagac gtcagggata tttataacaa acccccttc aagcaagtga tgctgaaggg 1560
ataacgggaa cgcagcggca ggatggaaga gacaggcact gcgctgcgga atgcctggga 1620
ggaaaagggg gagacctttc atccaggatg agggacattt aagatgaaat gtccgtggca 1680
ggatcgtttc tcttcactgc tgcatgcggc actgggaact cgccccacct gtgtccggaa 1740
cctgctcgct cacgtcggct ttcccttct gttttgttct aggacttctg cacggacctg 1800
gccgtctcca gtgccaactt cattcccacg gtcactgcca tctcgaccag tccggacctg 1860
cagtggctgg tgcagcccgc cctcgtctcc tctgtggccc catcgcagac cagagcccct 1920
caccctttcg gagtccccgc cccctccgct ggggcttact ccagggctgg cgttgtgaag 1980
accatgacag gaggccgagc gcagagcatt ggcaggaggg gcaaggtgga acaggtgagg 2040
aactctagcg tactcttcct gggaatgtgg gggctgggtg ggaagcagcc ccggagatgc 2100
aggagcccag tacagaggat gaagccactg atggggctgg ctgcacatcc gtaactggga 2160
gccctggctc caagcccatt ccatcccaac tcagactctg agtctcaccc taagaagtac 2220
tctcatagtt tcttccctaa gtttcttacc gcatgctttc agactgggct cttctttgtt 2280
ctcttgctga ggatcttatt ttaaatgcaa gtcacaccta ttctgcaact gcaggtcaga 2340
aatggtttca cagtggggtg ccaggaagca gggaagctgc aggagccagt tctactgggg 2400
tgggtgaatg gaggtgatgg cagacacttt tactgaatgt cggtcttttt ttgtgattat 2460
tctagttatc tccagaagaa gaagagaaaa ggagaatccg aagggaaagg aataagatgg 2520
ctgcagccaa atgccgcaac cggaggaggg agctgactga tacactccaa gcggtaggta 2580
ctctgtgggt tgctcctttt taaaacttaa gggaaagttg gagattgagc ataagggccc 2640
ttgagtaaga ctgtgtctta tgctttcctt tatccctctg tatacaggag acagaccaac 2700
tagaagatga gaagtctgct ttgcagaccg agattgccaa cctgctgaag gagaaggaaa 2760
aactagagtt catcctggca gctcaccgac ctgcctgcaa gatccctgat gacctgggct 2820
tcccagaaga gatgtctgtg gcttcccttg atctgactgg gggcctgcca gaggttgcca 2880
ccccggagtc tgaggaggcc ttcaccctgc ctctcctcaa tgaccctgag cccaagccct 2940
cagtggaacc tgtcaagagc atcagcagca tggagctgaa gaccgagccc tttgatgact 3000
tcctgttccc agcatcatcc aggcccagtg gctctgagac agcccgctcc gtgccagaca 3060
tggacctatc tgggtccttc tatgcagcag actgggagcc tctgcacagt ggctccctgg 3120
ggatggggcc catggccaca gagctggagc ccctgtgcac tccggtggtc acctgtactc 3180
ccagctgcac tgcttacacg tcttccttcg tcttcaccta ccccgaggct gactccttcc 3240

ccagctgtgc agctgcccac cgcaagggca gcagcagcaa tgagccttcc tctgactcgc 3300
tcagctcacc cacgctgctg gccctgtgag ggggcaggga aggggaggca gccggcaccc 3360
acaagtgcca ctgcccgagc tggtgcatta cagagaggag aaacacatct tccctagagg 3420
gttcctgtag acctagggag gaccttatct gtgcgtgaaa cacaccaggc tgtgggcctc 3480
aaggacttga aagcatccat gtgtggactc aagtccttac ctcttccgga gatgtagcaa 3540
aacgcatgga gtgtgtattg ttcccagtga cacttcagag agctggtagt tagtagcatg 3600
ttgagccagg cctgggtctg tgtctctttt ctctttctcc ttagtcttct catagcatta 3660
actaatctat tgggttcatt attggaatta acctggtgct ggatattttc aaattgtatc 3720
tagtgcagct gattttaaca ataactactg tgttcctggc aatagtgtgt tctgattaga 3780
aatgaccaat attatactaa gaaaagatac gactttattt tctggtagat agaaataaat 3840
agctatatcc atgtactgta gtttttcttc aacatcaatg ttcattgtaa tgttactgat 3900
catgcattgt tgaggtggtc tgaatgttct gacattaaca gttttccatg aaaacgtttt 3960
attgtgtttt taatttattt attaagatgg attctcagat atttatattt ttatttttatt 4020
tttttctacc ttgaggtctt ttgacatgtg gaaagtgaat ttgaatgaaa aatttaagca 4080
ttgtttgctt attgttccaa gacattgtca ataaaagcat ttaagttgaa tgcgaccaac 4140
cttgtgctct tttcattctg gaagtcttgt aagtttctga aaggtattat tggagaccag 4200
tttgtcaaga agggtagctg ctggaggggg acacaccctc tgtctgatcc cttatcaaag 4260
aggacaagga aactatagag ctgattttag aatattttac aaatacatgc cttccattgg 4320
aatgctaaga ttttctactg cttctgggga cgggaaaccg ctgtgtaaca gcttttgtgg 4380
gaatacattt tttctgtttc agtactcgca gggggaaata tttaaatttt gttgtgctaa 4440
tattaaattc agatgttttg atcttaaagg aacccttaa gcaaacagaa cctagctttg 4500
tacagactat tttaactttt tattctcaca aaatcacgtg gagggttatt ctacttcaaa 4560
gatgagcaaa ttgaagaatg gttagaataa acaactttct tgatattccg ttatcggcat 4620
tagaatcttc ctgctcgtta tcgtatccag caggctgaac tgcctcttga tacttggtta 4680
aaaaaaattt tcaggccggg cgcggtggcc catgcctgta atcctagcac tttgggaggc 4740
cgaggcaggc ggatcacctg aggtcgggag ttcgagacca gcctgaccaa catggagaaa 4800
ccccgtcttt actaaaaata caaaattagc ctggtgtggt ggtgcatgcc tgtaatccta 4860
gctacttgag aggctgagac aggaaaatca cttgaactcg ggaggcggat gttgcagcga 4920
actgagattg cgccattgca ctccagcctg ggcaacaaga ttgaaactct gtttaaaaaa 4980

aaaagttttc actaatgtgt acatttttt gtactctttt attctcgaaa gggaaggagg 5040

gctattgccc tatcccttat taataaatgc attgtggttt ctggtttctc taataccata 5100

tgcccttcat tcagtttata gtgggcggaa gtgggggaga aaaagttgct cagaaatcaa 5160

aagatatctc aaacagcaca aataatggct gatcgttctg caaacaaaaa gttacataat 5220

agctcaagaa ggagaagtca acatgactct gaacaagctt taacttagaa actttatcat 5280

cttaaggaag aacgtgacct ttgtccagga cgtctctggt aatggggcac ttacacacac 5340

atgcacacgt acaaaccaca gggaaaggag accgcccttc tgcctctgct cgcgagtatc 5400

acgcaggcac catgcactat gttttcacac acactgggtg gaagaagagc ttcagcgcca 5460

gtcttctaat gctttggtga taatgaaaat cactgggtgc ttatggggtg tcatattcaa 5520

tcgagttaaa agttttaatt caaaatgaca gttttactga ggttgatgtt ctcgtctatg 5580

atatctctgc ccctcccata aaaatggaca tttaaaagca acttaccgct ctttagatca 5640

ctcctatatc acacaccact tggggtgctg tttctgctag acttgtgatg acagtggcct 5700

taggatccct gtttgctgtt caaagggcaa atattttata gcctttaaat atacctaaac 5760

taaatacaga attaatataa ctaacaaaca cctggtctga aataacaagg tgatctaccc 5820

tggaaggaac ccagctggtg ggccaggagc ggtggctcac acctgtaatt ccagcacttt 5880

gggaggctga gacaggagga tcactggagt ccaggagttt gagaccagcc tgggcaacat 5940

ggcaaaaccc agtgtgcttc tgttgtccca gctacactac tcaggaggct gaggcaggag 6000

tatgacttga gcctgggagg gggaggttgc agagaactga tattgcacca ccactgcact 6060

ccagcctggg tgacacagca aaaccctatc tcaaaaaaaa aaaaaaaaaa aaggaaccca 6120

gctggttcct gtaggtgtgc aataataaca accagaggaa gaaaaggaag acgatttccc 6180

agatgaagaa gggcagctgg accttcggac 6210

<210> 4

<211> 2136

<212> DNA

<213> Homo sapiens

<400> 4

cgcgagccgc ctcctcccct tccccacgct cgaggagggg ggcgcggggg cccggctccg 60

gcgacggcca atcggagcgc acttccgtgg ctgactagcg cggtataaag gcgtgtggct 120

caggctgagc ggctgggacc ttgagagcgg ccaggccagc ctcggagcca gcagggagct 180

gggagctggg ggaaacgacg ccaggaaagc tatcgcgcca gagagggcga cgggggctcg 240
ggaagcctga cagggctttt gcgcacagct gccggctggc tgctacccgc ccgcgccagc 300
ccccgagaac gcgcgaccag gcacccagtc cggtcaccgc agcggagagc tcgccgctcg 360
ctgcagcgag gcccggagcg gccccgcagg gaccctcccc agaccgcctg ggccgcccgg 420
atgtgcacta aaatggaaca gcccttctac cacgacgact catacacagc tacgggatac 480
ggccgggccc ctggtggcct ctctctacac gactacaaac tcctgaaacc gagcctggcg 540
gtcaacctgg ccgaccccta ccggagtctc aaagcgcctg gggctcgcgg acccggccca 600
gagggcggcg gtggcggcag ctacttttct ggtcagggct cggacaccgg cgcgtctctc 660
aagctcgcct cttcggagct ggaacgcctg attgtcccca acagcaacgg cgtgatcacg 720
acgacgccta cacccccggg acagtacttt taccccgcg ggggtggcag cggtggaggt 780
gcaggggcg caggggcgg cgtcaccgag gagcaggagg gcttcgccga cggctttgtc 840
aaagccctgg acgatctgca caagatgaac cacgtgacac cccccaacgt gtccctgggc 900
gctaccgggg ggcccccggc tgggcccggg ggcgtctacg ccggcccgga gccacctccc 960
gtttacacca acctcagcag ctactcccca gcctctgcgt cctcgggagg cgccggggct 1020
gccgtcggga ccgggagctc gtacccgacg accaccatca gctacctccc acacgcgccg 1080
cccttcgccg gtggccaccc ggcgcagctg ggcttgggcc gcggcgcctc caccttcaag 1140
gaggaaccgc agaccgtgcc ggaggcgcgc agccgggacg ccacgccgcc ggtgtccccc 1200
atcaacatgg aagaccaaga gcgcatcaaa gtggagcgca agcggctgcg gaaccggctg 1260
gcggccacca agtgccggaa gcggaagctg gagcgcatcg cgcgcctgga ggacaaggtg 1320
aagacgctca aggccgagaa cgcggggctg tcgagtaccg ccggcctcct ccgggagcag 1380
gtggcccagc tcaaacagaa ggtcatgacc cacgtcagca acggctgtca gctgctgctt 1440
ggggtcaagg gacacgcctt ctgaacgtcc cctgcccctt tacggacacc ccctcgcttg 1500
gacggctggg cacacgcctc ccactggggt ccagggagca ggcggtgggc acccaccctg 1560
ggacctaggg gcgccgcaaa ccacactgga ctccggccct cctaccctgc gcccagtcct 1620
tccacctcga cgtttacaag ccccccccttc cacttttttt tgtatgtttt ttttctgctg 1680
gaaacagact cgattcatat tgaatataat atatttgtgt atttaacagg gaggggaaga 1740
gggggcgatc gcggcggagc tggccccgcc gcctggtact caagcccgcg gggacattgg 1800
gaaggggacc cccgcccccct gccctcccct ctctgcaccg tactgtggaa aagaaacacg 1860
cacttagtct ctaaagagtt tattttaaga cgtgtttgtg tttgtgtgtg tttgttcttt 1920

ttattgaatc tatttaagta aaaaaaaaat tggttcttta ttaatttctg ttgtcttttt 1980

ttccaagctg ggagggcggg gggaaaaaaa aagcactggt ttgcccccag ctcagtgctg 2040

ttggtggctc ggtcctgtat gtgtcccccc cgtcggttcg gcgcaggcat cttgtggtcc 2100

cagcccagga gtcccaccct tcccgcgtcc ccagat 2136

<210> 5

<211> 1398

<212> DNA

<213> Homo sapiens

<400> 5

tagatggcaa cctccctatc tgcccgcagg tcatagaggc gacacgtagc gtcatctgac 60

cctgaagcaa aggcatctcc actccaaagt tagacaaaat gccaggaatg ttcttctctg 120

ctaacccaaa ggaattgaaa ggaaccactc attcacttct agacgacaaa atgcaaaaaa 180

ggaggccaaa gacttttgga atggatatga aagcatacct gagatctatg atcccacatc 240

tggaatctgg aatgaaatct tccaagtcca aggatgtact ttctgctgct gaagtaatgc 300

aatggtctca atctctggaa aaacttcttg ccaaccaaac tggtcaaaat gtctttggaa 360

gtttcctaaa gtctgaattc agtgaggaga atattgagtt ctggctggct tgtgaagact 420

ataagaaaac agagtctgat cttttgccct gtaaagcaga agagatatat aaagcatttg 480

tgcattcaga tgctgctaaa caaatcaata ttgacttccg cactcgagaa tctacagcca 540

agaagattaa agcaccaacc cccacgtgtt ttgatgaagc acaaaaagtc atatatactc 600

ttatggaaaa ggactcttat cccaggttcc tcaaatcaga tatttactta aatcttctaa 660

atgacctgca ggctaatagc ctaaagtgac tggtccctgg ctgaagggaa ttaacagata 720

gtatcaaggc acgaaggaat gtgccagtat ggctccctgg gtgaacagct tggccttttt 780

tgggtgtctt gacaggccaa gaagaacaaa tgactcagaa tggattaaca tgaaagttat 840

ccaggcgcag agttgaagaa gcataagcaa gacaaaaaca gagagaccgc agaaggagga 900

agatactgtg gtactgtcat aaaaaacagt ggagctctgt attagaaagc ccctcagaac 960

tgggaaggcc aggtaactct agttacacag aaactgtgac taaagtctat gaaactgatt 1020

acaacaggct gtaagaatca aagtcaactg acatctatgc tacatattat tatatagttt 1080

gtactgagct attgaagtcc cattaactta aagtatatgt tttcaaattg ccattgctac 1140

tattgcttgt cggtgtattt tattttattg tttttgactt tggaagagat gaactgtgta 1200

tttaacttaa gctattgctc ttaaaaccag ggatcagaat atatttgtaa gttaaatcat 1260
tggtgctaat aataaatgtg gattttgtat taaaatatat agaagcaatt tctgtttaca 1320
tgtccttgct acttttaaaa acttgcattt attcctcaga ttttaaaaat aaataaataa 1380
ttcatttaaa aaaaaaaa 1398

<210> 6
<211> 2000
<212> DNA
<213> Homo sapiens

<400> 6
tttgggctgt gtgtgcgacg cgggtcggag gggcagtcgg gggaaccgcg aagaagccga 60
ggagcccgga gccccgcgtg acgctcctct ctcagtccaa aagcggcttt tggttcggcg 120
cagagagacc cgggggtcta gcttttcctc gaaaagcgcc gccctgccct tggccccgag 180
aacagacaaa gagcaccgca gggccgatca cgctgggggc gctgaggccg gccatggtca 240
tggaagtggg caccctggac gctggaggcc tgcgggcgct gctgggggag cgagcggcgc 300
aatgcctgct gctggactgc cgctccttct tcgctttcaa cgccggccac atcgccggct 360
ctgtcaacgt gcgcttcagc accatcgtgc ggcgccgggc caagggcgcc atgggcctgg 420
agcacatcgt gcccaacgcc gagctccgcg gccgcctgct ggccggcgcc taccacgccg 480
tggtgttgct ggacgagcgc agcgccgccc tggacggcgc caagcgcgac ggcaccctgg 540
ccctggcggc cggcgcgctc tgccgcgagg cgcgcgccgc gcaagtcttc ttcctcaaag 600
gaggatacga agcgttttcg gcttcctgcc cggagctgtg cagcaaacag tcgaccccca 660
tggggctcag ccttcccctg agtactagcg tccctgacag cgcggaatct gggtgcagtt 720
cctgcagtac cccactctac gatcagggtg gcccggtgga aatcctgccc tttctgtacc 780
tgggcagtgc gtatcacgct tcccgcaagg acatgctgga tgccttgggc ataactgcct 840
tgatcaacgt ctcagccaat tgtcccaacc attttgaggg tcactaccag tacaagagca 900
tccctgtgga ggacaaccac aaggcagaca tcagctcctg gttcaacgag gccattgact 960
tcatagactc catcaagaat gctggaggaa gggtgtttgt ccactgccag gcaggcattt 1020
cccggtcagc caccatctgc cttgcttacc ttatgaggac taatcgagtc aagctggacg 1080
aggcctttga gtttgtgaag cagaggcgaa gcatcatctc tcccaacttc agcttcatgg 1140

gccagctgct gcagtttgag tcccaggtgc tggctccgca ctgttcggca gaggctggga 1200
gccccgccat ggctgtgctc gaccgaggca cctccaccac caccgtgttc aacttccccg 1260
tctccatccc tgtccactcc acgaacagtg cgctgagcta ccttcagagc cccattacga 1320
cctctcccag ctgctgaaag gccacgggag gtgaggctct tcacatccca ttgggactcc 1380
atgctccttg agaggagaaa tgcaataact ctgggagggg ctcgagaggg ctggtcctta 1440
tttatttaac ttcacccgag ttcctctggg tttctaagca gttatggtga tgacttagcg 1500
tcaagacatt tgctgaactc agcacattcg ggaccaatat atagtgggta catcaagtcc 1560
atctgacaaa atggggcaga agagaaagga ctcagtgtgt gatccggttt ctttttgctc 1620
gcccctgttt tttgtagaat ctcttcatgc ttgacatacc taccagtatt attcccgacg 1680
acacatatac atatgagaat ataccttatt tatttttgtg taggtgtctg ccttcacaaa 1740
tgtcattgtc tactcctaga agaaccaaat acctcaattt ttgtttttga gtactgtact 1800
atcctgtaaa tatatcttaa gcaggtttgt tttcagcact gatggaaaat accagtgttg 1860
ggtttttttt tagttgccaa cagttgtatg tttgctgatt atttatgacc tgaaataata 1920
tatttcttct tctaagaaga cattttgtta cataaggatg actttttttat acaatggaat 1980
aaattatggc atttctattg 2000

<210> 7
<211> 1110
<212> DNA
<213> Homo sapiens

<400> 7
gacagagccc gggccacgga gctccttgcc agctctcctc ctcgcacagc cgctcgaacc 60
gcctgctgag ccccatggcc cgcgccacgc tctccgccgc ccccagcaat ccccggctcc 120
tgcgggtggc gctgctgctc ctgctcctgg tggccgccag ccggcgcgca gcaggagcgc 180
ccctggccac tgaactgcgc tgccagtgct tgcagaccct gcagggaatt cacctcaaga 240
acatccaaag tgtgaaggtg aagtcccccg gaccccactg cgcccaaacc gaagtcatag 300
ccacactcaa gaatgggcag aaagcttgtc tcaaccccgc atcgcccatg gttaagaaaa 360
tcatcgaaaa gatgctgaaa aatggcaaat ccaactgacc agaaggaagg aggaagctta 420
ttggtggctg ttcctgaagg aggccctgcc ttacaggaac agaagaggaa agagagacac 480
agctgcagag gccacctggc ttgcgcctaa tgtgtttgag catacttagg agaagtcttc 540

tatttattta tttatttatt tatttgtttg ttttagaaga ttctatgtta atattttatg 600
tgtaaaataa ggttatgatt gaatctactt gcacactctc ccattatatt tattgtttat 660
tttaggtcaa acccaagtta gttcaatcct gattcatatt taatttgaag atagaaggtt 720
tgcagatatt ctctagtcat ttgttaatat ttcttcgtga tgacatatca catgtcagcc 780
actgtgatag aggctgagga atccaagaaa atggccagta agatcaatgt gacggcaggg 840
aaatgtatgt gtgtctattt tgtaactgta aagatgaatg tcagttgtta tttattgaaa 900
tgatttcaca gtgtgtggtc aacatttctc atgttgaagc tttaagaact aaaatgttct 960
aaatatccct tggcatttta tgtctttctt gtaagatact gccttgttta atgttaatta 1020
tgcagtgttt ccctctgtgt tagagcagag aggtttcgat atttattgat gttttcacaa 1080
agaacaggaa aataaaatat ttaaaaatat 1110

<210> 8
<211> 1064
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (27)..(27)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (766)..(766)
<223> n=a, c, g or t

<400> 8
cacagccggg tcgcaggcac ctccccngcc agctctcccg cattctgcac agcttcccga 60
cgcgtctgct gagccccatg gcccacgcca cgctctccgc cgcccccagc aatccccggc 120
tcctgcgggt ggcgctgctg ctcctgctcc tggtgggcag ccggcgcgca gcaggagcgt 180

ccgtggtcac tgaactgcgc tgccagtgct tgcagacact gcagggaatt cacctcaaga 240
acatccaaag tgtgaatgta aggtcccccg gaccccactg cgcccaaacc gaagtcatag 300
ccacactcaa gaatgggaag aaagcttgtc tcaaccccgc atcccccatg gttcagaaaa 360
tcatcgaaaa gatactgaac aaggggagca ccaactgaca ggagagaagt aagaagctta 420
tcagcgtatc attgacactt cctgcagggt ggtccctgcc cttaccagag ctgaaaatga 480
aaaagagaac agcagctttc tagggacagc tggaaaggga cttaatgtgt ttgactattt 540
cttacgaggg ttctacttat ttatgtattt atttttgaaa gcttgtattt taatatttta 600
catgctgtta tttaaagatg tgagtgtgtt tcatcaaaca tagctcagtc ctgattattt 660
aattggaata tgatgggttt taaatgtgtc attaaactaa tatttagtgg gagaccataa 720
tgtgtcagcc accttgataa atgacagggt ggggaactgg agggtngggg gattgaaatg 780
caagcaatta gtggatcact gttagggtaa gggaatgtat gtacacatct atttttata 840
ctttttttt taaaaaagaa tgtcagttgt tatttattca aattatctca cattatgtgt 900
tcaacatttt tatgctgaag tttcccttag acattttatg tcttgcttgt agggcataat 960
gccttgttta atgtccattc tgcagcgttt ctctttccct tggaaaagag aatttatcat 1020
tactgttaca tttgtacaaa tgacatgata ataaaagttt tatg 1064

<210> 9
<211> 1172
<212> DNA
<213> Homo sapiens

<400> 9
gagacattcc tcaattgctt agacatattc tgagcctaca gcagaggaac ctccagtctc 60
agcaccatga atcaaactgc gattctgatt tgctgcctta tctttctgac tctaagtggc 120
attcaaggag tacctctctc tagaaccgta cgctgtacct gcatcagcat tagtaatcaa 180
cctgttaatc caaggtcttt agaaaaactt gaaattattc ctgcaagcca attttgtcca 240
cgtgttgaga tcattgctac aatgaaaaag aagggtgaga agagatgtct gaatccagaa 300
tcgaaggcca tcaagaattt actgaaagca gttagcaagg aaatgtctaa aagatctcct 360
taaaaccaga ggggagcaaa atcgatgcag tgcttccaag gatggaccac acagaggctg 420
cctctcccat cacttcccta catggagtat atgtcaagcc ataattgttc ttagtttgca 480
gttacactaa aaggtgacca atgatggtca ccaaatcagc tgctactact cctgtaggaa 540

ggttaatgtt catcatccta agctattcag taataactct accctggcac tataatgtaa 600
gctctactga ggtgctatgt tcttagtgga tgttctgacc ctgcttcaaa tatttccctc 660
acctttccca tcttccaagg gtactaagga atctttctgc tttggggttt atcagaattc 720
tcagaatctc aaataactaa aaggtatgca atcaaatctg ctttttaaag aatgctcttt 780
acttcatgga cttccactgc catcctccca aggggcccaa attctttcag tggctaccta 840
catacaattc caaacacata caggaaggta gaaatatctg aaaatgtatg tgtaagtatt 900
cttatttaat gaaagactgt acaaagtata agtcttagat gtatatattt cctatattgt 960
tttcagtgta catggaataa catgtaatta agtactatgt atcaatgagt aacaggaaaa 1020
ttttaaaaat acagatagat atatgctctg catgttacat aagataaatg tgctgaatgg 1080
ttttcaaata aaaatgaggt actctcctgg aaatattaag aaagactatc taaatgttga 1140
aagatcaaaa ggttaataaa gtaattataa ct 1172

<210> 10
<211> 4068
<212> DNA
<213> Homo sapiens

<400> 10
gaattcctcc tctcttcacc ccgttagctg ttttcaatgt aatgctgccg tccttctctt 60
gcactgcctt ctgcgctaac acctccattc ctgtttataa ccgtgtattt attacttaat 120
gtatataatg taatgttttg taagttatta atttatatat ctaacattgc ctgccaatgg 180
tggtgttaaa tttgtgtaga aaactctgcc taagagttac gacttttttct tgtaatgttt 240
tgtattgtgt attatataac ccaaacgtca cttagtagag acatatggcc cccttggcag 300
agaggacagg ggtgggcttt tgttcaaagg gtctgccctt tccctgcctg agttgctact 360
tctgcacaac ccctttatga accagttttc acccgaattt tgactgtttc atttagaaga 420
aaagcaaaat gagaaaaagc tttcctcatt tctccttgag atggcaaagc actcagaaat 480
gacatcacat accctaaaga accctgggat gactaaggca gagagagtct gagaaaactc 540
tttggtgctt ctgcctttag ttttaggaca catttatgca gatgagctta taagagaccg 600
ttccctccgc cttcttcctc agaggaagtt tcttggtaga tcaccgacac ctcatccagg 660
cgggggggttg gggggaaact tggcaccagc catcccaggc agagcaccac tgtgatttgt 720

tctcctggtg gagagagctg gaaggaagga gccagcgtgc aaataatgaa ggagcacggg 780
ggcaccttca gtagcaccgg aatcagcggt ggtagcggtg actctgctat ggacagcctg 840
cagccgctcc agcctaacta catgcctgtg tgtttgtttg cagaagaatc ttatcaaaaa 900
ttagcaatgg aaacgctgga ggaattagac tggtgtttag accagctaga gaccatacag 960
acctaccggt ctgtcagtga gatggcttct aacaagttca aaagaatgct gaaccgggag 1020
ctgacacacc tctcagagat gagccgatca gggaaccagg tgtctgaata catttcaaat 1080
actttcttag acaagcagaa tgatgtggag atcccatctc ctacccagaa agacagggag 1140
aaaaagaaaa agcagcagct catgacccag ataagtggag tgaagaaatt aatgcatagt 1200
tcaagcctaa acaatacaag catctcacgc tttggagtca acactgaaaa tgaagatcac 1260
ctggccaagg agctggaaga cctgaacaaa tggggtctta acatctttaa tgtggctgga 1320
tattctcaca atagacccct aacatgcatc atgtatgcta tattccagga aagagacctc 1380
ctaaagacat tcagaatctc atctgacaca tttataacct acatgatgac tttagaagac 1440
cattaccatt ctgacgtggc atatcacaac agcctgcacg ctgctgatgt agcccagtcg 1500
acccatgttc tcctttctac accagcatta gacgctgtct tcacagattt ggagatcctg 1560
gctgccattt ttgcagctgc catccatgac gttgatcatc ctggagtctc caatcagttt 1620
ctcatcaaca caaattcaga acttgctttg atgtataatg atgaatctgt gttggaaaat 1680
catcaccttg ctgtgggttt caaactgctg caagaagaac actgtgacat cttcatgaat 1740
ctcaccaaga agcagcgtca gacactcagg aagatggtta ttgacatggt gttagcaact 1800
gatatgtcta aacatatgag cctgctggca gacctgaaga caatggtaga aacgaagaaa 1860
gttacaagtt caggcgttct tctcctagac aactataccg atcgcattca ggtccttcgc 1920
aacatggtac actgtgcaga cctgagcaac cccaccaagt ccttggaatt gtatcggcaa 1980
tggacagacc gcatcatgga ggaatttttc cagcagggag acaaagagcg ggagagggga 2040
atggaaatta gcccaatgtg tgataaacac acagcttctg tggaaaaatc ccaggttggt 2100
ttcatcgact acattgtcca tccattgtgg gagacatggg cagatttggt acagcctgat 2160
gctcaggaca ttctcgatac cttagaagat aacaggaact ggtatcagag catgatacct 2220
caaagtccct caccaccact ggacgagcag aacagggact gccagggtct gatggagaag 2280
tttcagtttg aactgactct cgatgaggaa gattctgaag gacctgagaa ggagggagag 2340
ggacacagct atttcagcag cacaaagacg ctttgtgtga ttgatccaga aaacagagat 2400
tccctgggag agactgacat agacattgca acagaagaca agtcccccgt ggatacataa 2460
tccccctctc cctgtggaga tgaacattct atccttgatg agcatgccag ctatgtggta 2520

gggccagccc accatggggg ccaagacctg cacaggacaa gggccacctg gcctttcagt 2580
tacttgagtt tggagtcaga aagcaagacc aggaagcaaa tagcagctca ggaaatccca 2640
cggttgactt gccttgatgg caagcttggt ggagagggct gaagctgttg ctgggggccg 2700
attctgatca agacacatgg cttgaaaatg gaagacacaa aactgagaga tcattctgca 2760
ctaagtttcg ggaacttatc cccgacagtg actgaactca ctgactaata acttcattta 2820
tgaatcttct cacttgtccc tttgtctgcc aacctgtgtg ccttttttgt aaaacatttt 2880
catgtcttta aaatgcctgt tgaatacctg gagtttagta tcaacttcta cacagataag 2940
ctttcaaagt tgacaaactt ttttgactct ttctggaaaa gggaaagaaa atagtcttcc 3000
ttctttcttg ggcaatatcc ttcactttac tacagttact tttgcaaaca gacagaaagg 3060
atacacttct aaccacattt tacttccttc ccctgttgtc cagtccaact ccacagtcac 3120
tcttaaaact tctctctgtt tgcctgcctc caacagtact tttaactttt tgctgtaaac 3180
agaataaaat tgaacaaatt agggggtaga aaggagcagt ggtgtcgttc accgtgagag 3240
tctgcataga actcagcagt gtgccctgct gtgtcttgga ccctgccccc cacaggagtt 3300
gctacagtcc ctggccctgc ttcccatcct cctctcttca ccccgttagc tgttttcaat 3360
gtaatgctgc cgtccttctc ttgcactgcc ttctgcgcta acacctccat tcctgtttat 3420
aaccgtgtat ttattactta atgtatataa tgtaatgttt tgtaagttat taatttatat 3480
atctaacatt gcctgccaat ggtggtgtta aatttgtgta gaaaactctg cctaagagtt 3540
acgacttttt cttgtaatgt tttgtattgt gtattatata acccaaacgt cacttagtag 3600
agacatatgg ccccccttggc agagaggaca ggggtgggct tttgttcaaa gggtctgccc 3660
tttccctgcc tgagttgcta cttctgcaca acccctttat gaaccagttt tggaaacaat 3720
attctcacat tagatactaa atggtttata ctgagtcttt tacttttgta tagcttgata 3780
ggggcagggg caatgggatg tagtttttac ccaggttcta tccaaatcta tgtgggcatg 3840
agttgggtta taactggatc ctactatcat tgtggctttg gttcaaaagg aaacactaca 3900
tttgctcaca gatgattctt ctgattcttc tgaatgctcc cgaactactg actttgaaga 3960
ggtagcctcc tgcctgccat taagcaggaa tgtcatgttc cagttcatta caaaagaaaa 4020
caataaaaca atgtgaattt ttataataaa aaaaaaaaaa aggaattc 4068

<210> 11

<211> 2900

<212> DNA

<213> Homo sapiens

<400> 11
atgagtgtca gaagtgtgaa gggtgcctgt tctgaatccc agagcctcct ctccctctgt 60

gaggctggca ggtgaggaag ggtttaacct cactggaagg aatccctgga gctagcggct 120

gctgaaggcg tcgaggtgtg ggggcacttg gacagaacag tcaggcagcc gggagctctg 180

ccagctttgg tgaccttggg ccgggctggg agcgctgcgg cgggagccgg aggactatga 240

gctgccgcgc gttgtccaga gcccagccca gccctacgcg cgcggcccgg agctctgttc 300

cctggaactt tgggcactgc ctctgggacc cctgccggcc agcaggcagg atggtgcttg 360

cctcgtgccc cttggtgccc gtctgctgat gtgcccagcc tgtgcccgcc atgccgccct 420

ccatctcagc tttccaggcc gcctacatcg gcatcgaggt gctcatcgcc ctggtctctg 480

tgcccgggaa cgtgctggtg atctgggcgg tgaaggtgaa ccaggcgctg cgggatgcca 540

ccttctgctt catcgtgtcg ctggcggtgg ctgatgtggc cgtgggtgcc ctggtcatcc 600

ccctcgccat cctcatcaac attgggccac agacctactt ccacacctgc ctcatggttg 660

cctgtccggt cctcatcctc acccagagct ccatcctggc cctgctggca attgctgtgg 720

accgctacct ccgggtcaag atccctctcc ggtacaagat ggtggtgacc ccccggaggg 780

cggcggtggc catagccggc tgctggatcc tctccttcgt ggtgggactg acccctatgt 840

ttggctggaa caatctgagt gcggtggagc gggcctgggc agccaacggc agcatggggg 900

agcccgtgat caagtgcgag ttcgagaagg tcatcagcat ggagtacatg gtctacttca 960

acttctttgt gtgggtgctg ccccgcttc tcctcatggt cctcatctac ctggaggtct 1020

tctacctaat ccgcaagcag ctcaacaaga aggtgtcggc ctcctccggc gacccgcaga 1080

agtactatgg gaaggagctg aagatcgcca agtcgctggc cctcatcctc ttcctctttg 1140

ccctcagctg gctgcctttg cacatcctca actgcatcac cctcttctgc ccgtcctgcc 1200

acaagcccag catccttacc tacattgcca tcttcctcac gcacggcaac tcggccatga 1260

accccattgt ctatgccttc cgcatccaga agttccgcgt caccttcctt aagatttgga 1320

atgaccattt ccgctgccag cctgcacctc ccattgacga ggatctccca gaagagaggc 1380

ctgatgacta gaccccgcct tccgctccca ccagcccaca tccagtgggg tctcagtcca 1440

gtcctcacat gcccgctgtc ccaggggtct ccctgagcct gccccagctg ggctgttggc 1500

tgggggcatg ggggaggctc tgaagagata cccacagagt gtggtccctc cactaggagt 1560

taactaccct acacctctgg gccctgcagg aggcctggga gggcaagggt cctacggagg 1620

gaccaggtgt ctagaggcaa cagtgttctg agcccccacc tgcctgacca tcccatgagc 1680

agtccagcgc ttcagggctg ggcaggtcct ggggaggctg agactgcaga ggagccacct 1740

gggctgggag aaggtgcttg ggcttctgcg gtgaggcagg ggagtctgct tgtcttagat 1800

gttggtggtg cagccccagg accaagctta aggagaggag agcatctgct ctgagacgga 1860

tggaaggaga gaggttgagg atgcactggc ctgttctgta ggagagactg gccagaggca 1920

gctaaggggc aggaatcaag gagcctccgt tcccacctct gaggactctg gaccccaggc 1980

cataccaggt gctagggtgc ctgctctcct tgccctgggc cagcccagga ttgtacgtgg 2040

gagaggcaga aagggtaggt tcagtaatca tttctgatga tttgctggag tgctggctcc 2100

acgccctggg gagtgagctt ggtgcggtag gtgctggcct caaacagcca cgaggtggta 2160

gctctgagcc ctccttcttg ccctgagctt tccggggagg agcctggagt gtaattacct 2220

gtcatctggg ccaccagctc cactggcccc cgttgccggg cctggactgt cctaggtgac 2280

cccatctctg ctgcttctgg gcctgatgga gaggagaaca ctagacatgc caactcggga 2340

gcattctgcc tgcctgggaa cggggtggac gagggagtgt ctgtaaggac tcagtgttga 2400

ctgtaggcgc ccctggggtg ggtttagcag gctgcagcag gcagaggagg agtacccccc 2460

tgagagcatg tgggggaagg ccttgctgtc atgtgaatcc ctcaataccc ctagtatctg 2520

gctgggtttt caggggcttt ggaagctctg ttgcaggtgt ccgggggtct aggactttag 2580

ggatctggga tctggggaag gaccaaccca tgccctgcca agcctggagc ccctgtgttg 2640

gggggcaagg tgggggagcc tggagcccct gtgtgggagg gcgaggcggg ggagcctgga 2700

gcccctgtgt gggagggcga ggcgggggat cctggagccc ctgtgtcggg gggcgaggga 2760

ggggaggtgg ccgtcggttg accttctgaa catgagtgtc aactccagga cttgcttcca 2820

agcccttccc tctgttggaa attgggtgtg ccctggctcc caagggaggc ccatgtgact 2880

aataaaaaac tgtgaaccct 2900

<210> 12

<211> 1716

<212> DNA

<213> Homo sapiens

<400> 12
gccaactccg tggtggtctg ggtgaatatc caggccaaga ccacaggcta tgacacgcac 60
tgctacatct tgaacctggc cattgccgac ctgtgggttg tcctcaccat cccagtctgg 120
gtggtcagtc tcgtgcagca caaccagtgg cccatgggcg agctcacgtg caaagtcaca 180
cacctcatct tctccatcaa cctcttcagc agcattttct tcctcacgtg catgagcgtg 240
gaccgctacc tctccatcac ctacttcacc aacaccccca gcagcaggaa gaagatggta 300
cgccgtgtcg tctgcatcct ggtgtggctg ctggccttct gcgtgtctct gcctgacacc 360
tactacctga agaccgtcac gtctgcgtcc aacaatgaga cctactgccg gtccttctac 420
cccgagcaca gcatcaagga gtggctgatc ggcatggagc tggtctccgt tgtcttgggc 480
tttgccgttc ccttctccat tatcgctgtc ttctacttcc tgctggccag agccatctcg 540
gcgtccagtg accaggagaa gcacagcagc cggaagatca tcttctccta cgtggtggtc 600
ttccttgtct gctggttgcc ctaccacgtg gcggtgctgc tggacatctt ctccatcctg 660
cactacatcc ctttcacctg ccggctggag cacgccctct tcacggccct gcatgtcaca 720
cagtgcctgt cgctggtgca ctgctgcgtc aaccctgtcc tctacagctt catcaatcgc 780
aactacaggt acgagctgat gaaggccttc atcttcaagt actcggccaa aacagggctc 840
accaagctca tcgatgcctc cagagtctca gagacggagt actctgcctt ggagcagagc 900
accaaatgat ctgccctgga gaggctctgg gacgggttta cttgtttttg aacagggtga 960
tgggccctat ggttttctag agcaaagcaa agtagcttcg ggtcttgatg cttgagtaga 1020
gtgaagaggg gagcacgtgc cccctgcatc cattytctct ttctcttgat gacgcagctg 1080
tcatttggct gtgcgtgctg acagttttgc aacaggcaga gctgtgtcgc acagcagtgc 1140
tgtgcgtcag agccagctga ggacaggctt gcctggactt ctgtaagata ggattttctg 1200
tgtttcctga atttttata tggtgatttg tatttaaatt ttaagacttt attttctcac 1260
tattggtgta ccttataaat gtatttgaaa gttaaatata ttttaaatat tgtttgggag 1320
gcatagtgct gacatatatt cagagtgttg tagttttaag gttagcgtga ctttcagttt 1380
tgactaagga tgacactaat tgttagctgt tttgaaatta tatatatata aatatataaa 1440
tatatgccag tcttggctga aatgttttat ttaccatagt tttatatctg tgtggtgttt 1500
tgtaccggca cgggatatgg aacgaaaact gctttgtaat gcagtttgtg acattaatag 1560
tattgtaaag ttacatttta aaataaacaa aaaactgttc tggactgcaa atctgcacac 1620
acaacgaaca gttgcatttc agagagttct ctcaatttgt aagttatttt tttttaataa 1680

agatttttgt ttcctaaaaa aaaaaaaaaa aaaaaa 1716

<210> 13

<211> 3565

<212> DNA

<213> Homo sapiens

<400> 13
gcagccgggc ggccgcagaa gcgcccaggc ccgcgcgcca cccctctggc gccaccgtgg 60

ttgagcccgt gacgtttaca ctcattcata aaacgcttgt tataaaagca gtggctgcgg 120

cgcctcgtac tccaaccgca tctgcagcga gcaactgaga agccaagact gagccggcgg 180

ccgcggcgca gcgaacgagc agtgaccgtg ctcctaccca gctctgcttc acagcgccca 240

cctgtctccg cccctcggcc cctcgcccgg ctttgcctaa ccgccacgat gatgttctcg 300

ggcttcaacg cagactacga ggcgtcatcc tcccgctgca gcagcgcgtc cccggccggg 360

gatagcctct cttactacca ctcacccgca gactccttct ccagcatggg ctcgcctgtc 420

aacgcgcagg taaggctggc ttcccgtcgc cgcggggccg ggggcttggg gtcgcggagg 480

aggagacacc gggcgggacg ctccagtaga tgagtagggg gctcccttgt gcctggaggg 540

aggctgccgt ggccggagcg gtgccggctc gggggctcgg gacttgctct gagcgcacgc 600

acgcttgcca tagtaagaat tggttccccc ttcgggaggc aggttcgttc tgagcaacct 660

ctggtctgca ctccaggacg gatctctgac attagctgga gcagacgtgt cccaagcaca 720

aactcgctaa ctagagcctg gcttcttcgg ggaggtggca gaaagcggca atccccctc 780

ccccggcagc ctggagcacg gaggagggat gagggaggag ggtgcagcgg gcgggtgtgt 840

aaggcagttt cattgataaa aagcgagttc attctggaga ctccggagcg gcgcctgcgt 900

cagcgcagac gtcagggata tttataacaa acccccttc aagcaagtga tgctgaaggg 960

ataacgggaa cgcagcggca ggatggaaga gacaggcact gcgctgcgga atgcctggga 1020

ggaaaagggg gagacctttc atccaggatg agggacattt aagatgaaat gtccgtggca 1080

ggatcgtttc tcttcactgc tgcatgcggc actgggaact cgccccacct gtgtccggaa 1140

cctgctcgct cacgtcggct ttcccttct gttttgttct aggacttctg cacggacctg 1200

gccgtctcca gtgccaactt cattcccacg gtcactgcca tctcgaccag tccggacctg 1260

cagtggctgg tgcagcccgc cctcgtctcc tctgtggccc catcgcagac cagagccct 1320

caccctttcg gagtccccgc cccctccgct ggggcttact ccagggctgg cgttgtgaag 1380

accatgacag gaggccgagc gcagagcatt ggcaggaggg gcaaggtgga acaggtgagg 1440
aactctagcg tactcttcct gggaatgtgg gggctgggtg ggaagcagcc ccggagatgc 1500
aggagcccag tacagaggat gaagccactg atggggctgg ctgcacatcc gtaactggga 1560
gccctggctc caagcccatt ccatcccaac tcagactctg agtctcaccc taagaagtac 1620
tctcatagtt tcttccctaa gtttcttacc gcatgctttc agactgggct cttctttgtt 1680
ctcttgctga ggatcttatt ttaaatgcaa gtcacaccta ttctgcaact gcaggtcaga 1740
aatggtttca cagtggggtg ccaggaagca gggaagctgc aggagccagt tctactgggg 1800
tgggtgaatg gaggtgatgg cagacacttt tactgaatgt cggtcttttt ttgtgattat 1860
tctagttatc tccagaagaa gaagagaaaa ggagaatccg aagggaaagg aataagatgg 1920
ctgcagccaa atgccgcaac cggaggaggg agctgactga tacactccaa gcggtaggta 1980
ctctgtgggt tgctcctttt taaaacttaa gggaaagttg gagattgagc ataagggccc 2040
ttgagtaaga ctgtgtctta tgctttcctt tatccctctg tatacaggag acagaccaac 2100
tagaagatga gaagtctgct ttgcagaccg agattgccaa cctgctgaag gagaaggaaa 2160
aactagagtt catcctggca gctcaccgac ctgcctgcaa gatccctgat gacctgggct 2220
tcccagaaga gatgtctgtg gcttcccttg atctgactgg gggcctgcca gaggttgcca 2280
ccccggagtc tgaggaggcc ttcaccctgc ctctcctcaa tgaccctgag cccaagccct 2340
cagtggaacc tgtcaagagc atcagcagca tggagctgaa gaccgagccc tttgatgact 2400
tcctgttccc agcatcatcc aggcccagtg gctctgagac agcccgctcc gtgccagaca 2460
tggacctatc tgggtccttc tatgcagcag actgggagcc tctgcacagt ggctccctgg 2520
ggatggggcc catggccaca gagctggagc ccctgtgcac tccggtggtc acctgtactc 2580
ccagctgcac tgcttacacg tcttccttcg tcttcaccta ccccgaggct gactccttcc 2640
ccagctgtgc agctgcccac cgcaagggca gcagcagcaa tgagccttcc tctgactcgc 2700
tcagctcacc cacgctgctg gccctgtgag ggggcaggga aggggaggca gccggcaccc 2760
acaagtgcca ctgcccgagc tggtgcatta cagagaggag aaacacatct tccctagagg 2820
gttcctgtag acctagggag gaccttatct gtgcgtgaaa cacaccaggc tgtgggcctc 2880
aaggacttga aagcatccat gtgtggactc aagtccttac ctcttccgga gatgtagcaa 2940
aacgcatgga gtgtgtattg ttcccagtga cacttcagag agctggtagt tagtagcatg 3000
ttgagccagg cctgggtctg tgtctctttt ctctttctcc ttagtcttct catagcatta 3060
actaatctat tgggttcatt attggaatta acctggtgct ggatattttc aaattgtatc 3120

tagtgcagct gattttaaca ataactactg tgttcctggc aatagtgtgt tctgattaga 3180
aatgaccaat attatactaa gaaaagatac gactttattt tctggtagat agaaataaat 3240
agctatatcc atgtactgta gtttttcttc aacatcaatg ttcattgtaa tgttactgat 3300
catgcattgt tgaggtggtc tgaatgttct gacattaaca gttttccatg aaaacgtttt 3360
attgtgtttt taatttattt attaagatgg attctcagat atttatattt ttatttttatt 3420
tttttctacc ttgaggtctt ttgacatgtg gaaagtgaat ttgaatgaaa aatttaagca 3480
ttgtttgctt attgttccaa gacattgtca ataaaagcat ttaagttgaa tgcgaccaac 3540
cttgtgctct tttcattctg gaagt 3565

<210> 14
<211> 982
<212> DNA
<213> Homo sapiens

<400> 14
agacaatttt cagcaggaag aagtagaaag gataaaatgg atcctgcacc acgggaacct 60
cacagcacct cacttctttt ggttttcttt ctctttgggg cacctctgga ctcactcccc 120
agcatgaagg cgctgagccc ggtgcgcggc tgctacgagg cggtgtgctg cctgtcggaa 180
cgcagtctgg ccatcgcccg gggccgaggg aagggcccgg cagctgagga gccgctgagc 240
ttgctggacg acatgaacca ctgctactcc cgcctgcggg aactggtacc cggagtcccg 300
agaggcactc agcttagcca ggtggaaatc ctacagcgcg tcatcgacta cattctcgac 360
ctgcaggtag tcctggccga gccagcccct ggacccctg atggccccca ccttcccatc 420
cagacagccg agctcgctcc ggaacttgtc atctccaacg acaaaaggag cttttgccac 480
tgactcggcc gtgtcctgac acctccagaa cgcaggtgct ggcgcccgtt ctgcctggga 540
ccccgggaac ctctcctgcc ggaagccgga cggcagggat gggccccaac ttcgccctgc 600
ccacttgact tcaccaaatc ccttcctgga gactgaacct ggtgctcagg agcgaaggac 660
tgtgaacttg tggcctgaag agccagagct agctctggcc accagctggg cgacgtcacc 720
ctgctcccac cccaccccaa gttctaaggt cttttcagag cgtggaggtg tggaaggagt 780
ggctgctctc caaactatgc caaggcggcg gcagagctgg tcttctggtc tccttggaga 840
aaggttctgt tgccctgatt tatgaactct ataatagagt atataggttt tgtacctttt 900
ttacaggaag gtgactttct gtaacaatgc gatgtatatt aaacttttta taaaagttaa 960

cattttgcat aataaacgat tt 982

<210> 15

<211> 3915

<212> DNA

<213> Homo sapiens

<400> 15
gtggggtggg gtggggctgg gggcttgtcg ccctttcagg ctccaccctt tgcggagatt 60
ataaatagtc atgatcccag cgagacccag agatgcctgt aatggtgaga ctttggatcc 120
ttcctgagga cgtggagaaa actttctgct gagaaggaca ttttgaaggt tttgttggct 180
gaaaaagctg tttctggaat cacccctaga tctttcttga agacttgaat tagattacag 240
cgatggggac acagaaggtc accccagctc tgatatttgc catcacagtt gctacaatcg 300
gctctttcca atttggctac aacactgggg tcatcaatgc tcctgagaag atcataaagg 360
aatttatcaa taaaactttg acggacaagg gaaatgcccc accctctgag gtgctgctca 420
cgtctctctg gtccttgtct gtggccatat tttccgtcgg gggtatgatc ggctcctttt 480
ccgtcggact cttcgtcaac cgctttggca ggcgcaattc aatgctgatt gtcaacctgt 540
tggctgtcac tggtggctgc tttatgggac tgtgtaaagt agctaagtcg gttgaaatgc 600
tgatcctggg tcgcttggtt attggcctct tctgcggact ctgcacaggt tttgtgccca 660
tgtacattgg agagatctcg cctactgccc tgcggggtgc ctttggcact ctcaaccagc 720
tgggcatcgt tgttggaatt ctggtggccc agatctttgg tctggaattc atccttgggt 780
ctgaagagct atggccgctg ctactgggtt ttaccatcct tcctgctatc ctacaaagtg 840
cagcccttcc attttgccct gaaagtccca gatttttgct cattaacaga aaagaagagg 900
agaatgctaa gcagatcctc cagcggttgt ggggcaccca ggatgtatcc caagacatcc 960
aggagatgaa agatgagagt gcaaggatgt cacaagaaaa gcaagtcacc gtgctagagc 1020
tctttagagt gtccagctac cgacagccca tcatcatttc cattgtgctc cagctctctc 1080
agcagctctc tgggatcaat gctgtgttct attactcaac aggaatcttc aaggatgcag 1140
gtgttcaaga gcccatctat gccaccatcg gcgcgggtgt ggttaatact atcttcactg 1200
tagtttctct atttctggtg gaaagggcag gaagaaggac tctgcatatg ataggccttg 1260
gagggatggc tttttgttcc acgctcatga ctgtttcttt gttattaaag gataactata 1320

atgggatgag ctttgtctgt attggggcta tcttggtctt tgtagccttc tttgaaattg 1380
gaccaggccc cattccctgg tttattgtgg ccgaactctt cagccagggc ccccgcccag 1440
ctgcgatggc agtggccggc tgctccaact ggacctccaa cttcctagtc ggattgctct 1500
tcccctccgc tgctcactat ttaggagcct acgtttttat tatcttcacc ggcttcctca 1560
ttaccttctt ggcttttacc ttcttcaaag tccctgagac ccgtggcagg acttttgagg 1620
atatcacacg ggcctttgaa gggcaggcac acggtgcaga tagatctgga aaggacggcg 1680
tcatggagat gaacagcatc gagcctgcta aggagaccac caccaatgtc taagtcgtgc 1740
ctccttccac ctccctcccg gcatgggaaa gccacctctc cctcaacaag ggagagacct 1800
catcaggatg aacccaggac gcttctgaat gctgctactt aattcctttc tcatcccacg 1860
cactccatga gcaccccaag gctgcggttt gttggatctt caatggcttt ttaaatttta 1920
tttcctggac atcctcttct gcttaggaga gaccgagtga acctaccttc atttcaggag 1980
ggattggccg cttggcacat gacaactttg ccagcttttc ctcccttggg ttctgatatt 2040
gccgcactag gggatatagg agaggaaaag taaggtgcag ttcccccaac ctcagactta 2100
ccaggaagca gatacatatg agtgtggaag ccggagggtg tttatgtaag agcaccttcc 2160
tcacttccat acagctctac gtggcaaatt aacttgagtt ttatttattt tatcctctgg 2220
tttaattaca taatttttt ttttttactt taagtttcag gatacatgtg ccgaatgtgc 2280
aggtttgtta cataggtata tatatgccat gatggaaata tttatttttt taagcgtaat 2340
tttgccaaat aataaaaaca gaaggaaatt gagattagag ggaggtgttt aaagagaggt 2400
tatagagtag aagatttgat gctggagagg ttaaggtgca ataagaattt agggagaaat 2460
gttgttcatt attggagggt aaatgatgtg gtgcctgagg tctgtacgtt acctcttaac 2520
aatttctgtc cttcagatgg aaactcttta acttctcgta aaagtcatat acctatataa 2580
taaagctact gatttccttg gagcttttt ctttaagata atagtttaca tgtagtagta 2640
cttgaaatct aggattatta actaatatgg gcattgtagt taatgatggt tgatgggttc 2700
taattttgga tggagtccag ggaagagaaa gtgatttcta gaaagcctgt tcccctcact 2760
ggatgaaata actccttctt gtagtagtct cattactttt gaagtaatcc cgccacctat 2820
ctcgtgggag agccatccaa ataagaaacc taaaataatt ggttcttggt agagattcat 2880
tatttttcca ctttgttctt taggagattt taggtgttga ttttctgttg tattttaact 2940
catacctta aaggaattcc ccaaagaatg tttatagcaa acttggaatt tgtaacctca 3000
gctctgggag aggatttttt tctgagcgat tattatctaa agtgtgttgt tgctttaggc 3060
tcacggcacg cttgcgtatg tctgttacca tgtcactgtg gtcctatgcc gaatgccctc 3120

aggggacttg aatctttcca ataaaccagg tttagacagt atgagtcaat gtgcagtgta 3180
gcccacactt gagaggatga atgtatgtgc actgtcactt tgctctgggt ggaagtacgt 3240
tattgttgac ttatttttctc tgtgtttgtt cctacagccc ctttttcata tgttgctcag 3300
tctccctttc ccttcttggt gcttacacat ctcagaccct ttagccaaac ccttgtcagt 3360
gacagtattt tggttcttag ttctcactgt tccctctgct cctggagcct ttgaataaaa 3420
atgcacgtag ctgaggccgg atgcggtggc tcacgcctgt aatcccagca ctttgggagg 3480
cctaggcggg cggtcagggg ttcgagacca gtctggccaa catcgtgaaa ccctgtctct 3540
actaaaaatg caaaaattag ccgggcgtgg tggcgggcgc ctgtaatccc agctacttgg 3600
gaagctgagg cgggagaatc atgtgaaccc gggacgcagg ggttgcagtg agcggagatc 3660
gcatcattgc actctagcct gggccacagg gcgagactcc gtctcaaaaa aaaaaaaatg 3720
cacatagcta tcgagtgtgc tttagcttga aaaggtgacc ttgcaacttc atgtcaactt 3780
tctggctcct caaacagtag gttggcagta aggcagggtc ccatttctca ctgagaagat 3840
tgtgaatatt tccatatgga ttttctattg ttactctggt tctttgtttt aaaataaaaa 3900
ttctgaatgt acacg 3915

<210> 16
<211> 1670
<212> DNA
<213> Homo sapiens

<400> 16
cggcacgagg caagtgacgc cgagggcctg agtgctccag tagccaccgc atctggagaa 60
ccagcggtta ccatggaggg gatcagtata tacacttcag ataactacac cgaggaaatg 120
ggctcagggg actatgactc catgaaggaa ccctgtttcc gtgaagaaaa tgctaatttc 180
aataaaatct tcctgcccac catctactcc atcatcttct taactggcat tgtgggcaat 240
ggattggtca tcctggtcat gggttaccag aagaaactga gaagcatgac ggacaagtac 300
aggctgcacc tgtcagtggc cgacctcctc tttgtcatca cgcttccctt ctgggcagtt 360
gatgccgtgg caaactggta ctttgggaac ttcctatgca aggcagtcca tgtcatctac 420
acagtcaacc tctacagcag tgtcctcatc ctggccttca tcagtctgga ccgctacctg 480
gccatcgtcc acgccaccaa cagtcagagg ccaaggaagc tgttggctga aaaggtggtc 540

tatgttggcg tctggatccc tgccctcctg ctgactattc ccgacttcat ctttgccaac 600
gtcagtgagg cagatgacag atatatctgt gaccgcttct accccaatga cttgtgggtg 660
gttgtgttcc agtttcagca catcatggtt ggccttatcc tgcctggtat tgtcatcctg 720
tcctgctatt gcattatcat ctccaagctg tcacactcca agggccacca gaagcgcaag 780
gccctcaaga ccacagtcat cctcatcctg gctttcttcg cctgttggct gccttactac 840
attgggatca gcatcgactc cttcatcctc ctggaaatca tcaagcaagg gtgtgagttt 900
gagaacactg tgcacaagtg gatttccatc accgaggccc tagctttctt ccactgttgt 960
ctgaacccca tcctctatgc tttccttgga gccaaattta aaacctctgc ccagcacgca 1020
ctcacctctg tgagcagagg gtccagcctc aagatcctct ccaaaggaaa gcgaggtgga 1080
cattcatctg tttccactga gtctgagtct tcaagttttc actccagcta acacagatgt 1140
aagagacttt tttttatacg ataaataact tttttttaag ttacacattt ttcagatata 1200
aaagactgac caatattgta cagtttttat tgcttgttgg atttttgctc ttgtgtttct 1260
ttagtttttc gtgaaggttt aattgactta tttatataaa tttttttttgt ttcatattga 1320
tgtgtgtcta ggcaggacct gtggccaagt tcttagttgc tgtatgtctc gtggtaggac 1380
tgtagaaaag ggaactgaac attccagagc gtgtagtgaa tcacgtaaag ctagaaatga 1440
tccccagctg tttatgcata gataatctct ccattcccgt ggaacgtttt tcctgttctt 1500
aagacgtgat tttgctgtag aagatggcac ttataaccaa agcccaaagt ggtatagaaa 1560
tgctggtttt tcagttttca ggagtgggtt gatttcagca cctacagtgt acagtcttgt 1620
attaagttgt taataaaagt acatgttaaa cttaaaaaaa aaaaaaaaaa 1670

<210> 17

<211> 2360

<212> DNA

<213> Homo sapiens

<400> 17
gctacgcggg ccacgctgct ggctggcctg acctaggcgc gcggggtcgg gcggccgcgc 60
gggcgggctg agtgagcaag acaagacact caagaagagc gagctgcgcc tgggtcccgg 120
ccaggcttgc acgcagaggc gggcggcaga cggtgcccgg cggaatctcc tgagctccgc 180
cgcccagctc tggtgccagc gcccagtggc cgccgcttcg aaagtgactg gtgcctcgcc 240
gcctcctctc ggtgcgggac catgaagctg ctgccgtcgg tggtgctgaa gctctttctg 300

gctgcagttc tctcggcact ggtgactggc gagagcctgg agcggcttcg gagagggcta 360
gctgctggaa ccagcaaccc ggaccctccc actgtatcca cggaccagct gctaccccta 420
ggaggcggcc gggaccggaa agtccgtgac ttgcaagagg cagatctgga ccttttgaga 480
gtcactttat cctccaagcc acaagcactg gccacaccaa acaaggagga gcacgggaaa 540
agaaagaaga aaggcaaggg gctagggaag aagagggacc catgtcttcg gaaatacaag 600
gacttctgca tccatggaga atgcaaatat gtgaaggagc tccgggctcc ctcctgcatc 660
tgccacccgg gttaccatgg agagaggtgt catgggctga gcctcccagt ggaaaatcgc 720
ttatatacct atgaccacac aaccatcctg gccgtggtgg ctgtggtgct gtcatctgtc 780
tgtctgctgg tcatcgtggg gcttctcatg tttaggtacc ataggagagg aggttatgat 840
gtggaaaatg aagagaaagt gaagttgggc atgactaatt cccactgaga gagacttgtg 900
ctcaaggaat cggctgggga ctgctacctc tgagaagaca caaggtgatt tcagactgca 960
gaggggaaag acttccatct agtcacaaag actccttcgt ccccagttgc cgtctaggat 1020
tgggcctccc ataattgctt tgccaaaata ccagagcctt caagtgccaa acagagtatg 1080
tccgatggta tctgggtaag aagaaagcaa aagcaaggga ccttcatgcc cttctgattc 1140
ccctccacca aaccccactt cccctcataa gtttgtttaa acacttatct tctggattag 1200
aatgccggtt aaattccata tgctccagga tctttgactg aaaaaaaaaa agaagaagaa 1260
gaaggagagc aagaaggaaa gatttgtgaa ctggaagaaa gcaacaaaga ttgagaagcc 1320
atgtactcaa gtaccaccaa gggatctgcc attgggaccc tccagtgctg gatttgatga 1380
gttaactgtg aaataccaca agcctgagaa ctgaattttg ggacttctac ccagatggaa 1440
aaataacaac tatttttgtt gttgttgttt gtaaatgcct cttaaattat atatttattt 1500
tattctatgt atgttaattt atttagtttt taacaatcta acaataatat ttcaagtgcc 1560
tagactgtta ctttggcaat ttcctggccc tccactcctc atccccacaa tctggcttag 1620
tgccacccac ctttgccaca aagctaggat ggttctgtga cccatctgta gtaatttatt 1680
gtctgtctac atttctgcag atcttccgtg gtcagagtgc cactgcggga gctctgtatg 1740
gtcaggatgt aggggttaac ttggtcagag ccactctatg agttggactt cagtcttgcc 1800
taggcgattt tgtctaccat ttgtgttttg aaagcccaag gtgctgatgt caaagtgtaa 1860
cagatatcag tgtctccccg tgtcctctcc ctgccaagtc tcagaagagg ttgggcttcc 1920
atgcctgtag ctttcctggt ccctcacccc catggcccca ggccacagcg tgggaactca 1980
ctttcccttg tgtcaagaca tttctctaac tcctgccatt cttctggtgc tactccatgc 2040

aggggtcagt gcagcagagg acagtctgga gaaggtatta gcaaagcaaa aggctgagaa 2100

ggaacaggga acattggagc tgactgttct tggtaactga ttacctgcca attgctaccg 2160

agaaggttgg aggtggggaa ggctttgtat aatcccaccc acctcaccaa aacgatgaag 2220

gtatgctgtc atggtccttt ctggaagttt ctggtgccat ttctgaactg ttacaacttg 2280

tatttccaaa cctggttcat atttatactt tgcaatccaa ataaagataa cccttattcc 2340

ataaaaaaaa aaaaaaaaaa 2360

<210> 18

<211> 6082

<212> DNA

<213> Homo sapiens

<400> 18

caccttgaca cataggccac ttcctgagcc tggacccagt ctcagagctg gggaactggc 60

ccaatgcaaa agggtcggga gcatctgcaa cagagactga gctctatcag cttcggtgac 120

atagcctcca ttcacgctcc ccaactcagc agagagagca caccatcaga cttctaagac 180

ttagtagcca agaagtgttg aattaaactc tctgagacct ctctttagtc tgaccctggc 240

agcctcagtc tcccagagcc tgtgggaact cggcagccga gaggcagaag gctgggcgac 300

gtccggagaa gaagaaacgg gggaagaact tttctcttac gatctggctt tactctcacg 360

cgcacagccg agtccctggg gacccagcag aggtccgaag cggagcgggg cggggcgggg 420

ctacggaagc tggcgaggcc gagcccctcc tagtggttcc ggaccttgct ccctgaacac 480

tcggaggtgg cggtggatct tactccttcc agccagtgag gatccagcaa cctgctccgt 540

gcctcccgcg cctgttggtt ggaagtgacg accttgaaga tcggccggtt ggaagtgacg 600

accttgaaga tcggcgggcg cagcggggcc gaggggggcgg gtctggcgct aggtccagcc 660

cctgcgtgcc gggaacccca gaggaggtcg cagttcagcc cagctgaggc ctgtctgcag 720

aatcgacacc aaccagcatc atgtccatga cactggggta ctgggacatc cgcggggtga 780

gtgagggtcc gctgcactgt gggaccgggc gcgtgggcgg gaagtgccga gcggctgggg 840

accggctcta gggacggttc cctccttagg gctatctctc acaggagggc ctgtgcatgc 900

ctgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gcgtgcgccg gggtgggggg 960

ggggtgcagt gcagtgtaga ctaggggctc acctggtgca gagaaagtca ccaagtcagg 1020

gaccctccat ctctgacacg acctgcgggc catctcttcc agctggccca cgccatgcgc 1080

ctgctcctgg aatacacaga ctcaagctac gaggaaaaga agtatacgat gggggacggt 1140
aatgacaccc ttgtgtccgg gctctgccca ctcacgctga gttggcacca agcaacccat 1200
ggtggccacc tgtggctacc tctgcaggcc tcccctgctg gagctgcagg ctgtcccttc 1260
cctgagcccc ggtgagggag tcctgtggcc ttgcaaggca gaatgctggg gcgggatgct 1320
gggccccctg tttaattggg ttgggtgtcc ctcagagctt ccctaaaccc tggaagcctt 1380
agccgtgtgg ggtccagagc cctcagcggg attctttgtc cctgaaccct gggatgtggg 1440
actgagtggt cagattctag atccacctgt ctcagggatc ttgccactgg ttccttggga 1500
gggtccccgg gaaggagggc tgggctctgg ggaggtttgt tttcacttct tcttccccac 1560
ggcagctcct gactatgaca gaagccagtg gctgaatgaa aaattcaagc tgggcctgga 1620
ctttcccaat gtaggtgcag gggaaggggc ggttttgggg gaaagtgcga cgtgtctctg 1680
actgcatctc ctctccccag attagaggtg ttcggatcag gagtcttctg cccaattcct 1740
ctcactcctg gttgtctaca cagcccctgc atgatgttct gtgtcccagc tcatttgttc 1800
atgtgacagt atttctatgt caggcctgcc atgagcgggc acagtgagtg cctggtctcc 1860
cctctgccct tgcatatggg aaggggatgc tggggagcct gctggcccca actgagcttc 1920
cccggtttcc catctatcca gctgccctac ttgattgatg gggctcacaa gatcacccag 1980
agcaacgcca tcctgtgcta cattgcccgc aagcacaacc tgtgtgagtg tggttggctg 2040
cagtgtgtgg ggggaaggtg gcatcctcct tggctggatt ggggtgctat gctcagagtg 2100
agtctgtgtt ttgtgggtgg caggtgggga gacagaagag gagaagattc gtgtggacat 2160
tttggagaac caggctatgg acgtctccaa tcagctggcc agagtctgct acagccctga 2220
ctttgtgagt ccctccctgg tctggaccag aagccaggct tgtattccca tctactctgg 2280
tcctattcac aatttgaact cctcactgct attgatcctc tgagggttcc ctgtactgta 2340
gcagagtgac tgtcatatca ttaaacttat aaaataaaaa cttgaaatca gtccccacca 2400
atcataggaa gtcctatgaa agctagcaat tcagttccta gacaataaag tcatgcgttc 2460
agaattccct tcacctctga cccctgacct ctgccacggg ccatctaacc cagctggttc 2520
atgccatctg cctgttcggg gaatacgcag actcacacta tggggtaaag aagtatgtag 2580
tgggatgtaa tagcaccctc ctctgtggaa ttccgtgcta ttcttttcct agtgcttctc 2640
cttgacacgc acagggatga atgcattttc ctaacaggca gctcaggggc ggccagaggg 2700
cctttgttcc tctgtctcct tccatgcagc ctctcaaaat ctcatctctc cagaaactac 2760
tcaatgtcca ttgtattatt ctgccactgc actaggagga actttcaact tctttctctg 2820

agctccttta gttctttgta tccttgattt tgctcgtgtc tgggtccaga gcctgccagg 2880
tactcaggtg ctcctgggct gacccagagg agggtggttg ggaggtcagt ggggacagat 2940
tcagggatag tgttgcattc ctctctgcct tcccatcacc acaaaagcct ccagctaccc 3000
atttggagtg taataaatgc tggtatgtcc agctgaagcc agttccagct gtggggaaga 3060
tggctgcttg ctcgtggcca gctggggcca tacacagccc tggggaggcc acatctgtgc 3120
agggagcttg tgtctgaggg tggtgacagc tgttttctgc ctcaggagaa actgaagcca 3180
gaatacttgg aggaacttcc tacaatgatg cagcacttct cacagttcct ggggaagagg 3240
ccatggtttg ttggagacaa ggtaatgggg gcatgtgatg aggacactag agatttgcca 3300
tacatcctat gttacagaga ttccagccca cacattcttg gccttctgca gatcaccttt 3360
gtagatttcc tcgcctatga tgtccttgac ctccaccgta tatttgagcc caactgcttg 3420
gacgccttcc caaatctgaa ggacttcatc tcccgctttg aggtgatgcc cccatcctcc 3480
tttctctttg atgccccttg ttccgttacc tcctttcaga tgctttccca gtcctggatc 3540
tgcataaaga ataacttgca tttattgagt gctggcttca tgccaggaac cttgcccagc 3600
acattatacc tatcgtgtgg aatttgaaat ttccaacatt cctacagggt gacagaatta 3660
tcttgcccat ttagagataa gaaaactttg aatgagaggg tcagtccttt gttctgggtc 3720
ccagagccag tggaggctgt gctgggctcc ctgtgagcat ctggatctat gggtggcagt 3780
cagggctctc ccttttgtga caaaagaaag aagcctcagg ctcatccagc ctggatttca 3840
cagcccaggg cactttggaa gaggcagaga actttaggag catggatgca gctggcaata 3900
gtaggactga cacacggtgg cattgacgtc gagtacgaaa cccacaggca gtattcatag 3960
ctactcccag aagctttgca cgatcagacc cccacgtggg gaatcctgag agccagagct 4020
gtggccagag ctggattagg gtacatatgt gggtgcccct gttgaaggag tgtatgttga 4080
agtgctctgt gctggggcac tctccttctt tatctttttt cctctctttt ttccctccag 4140
tgttccaagt gttccccctg tgagatgagt agcacactga ttttactgct attcaccgac 4200
cttctcctct gcatgaggca gggtgtgagg cacagtggga gttgcataga tgactgcccc 4260
atcctggaaa tgagtgcagt gagaggcctg caggcagagc agcctgtgag gtgtgtggca 4320
ccacctgggt accaggcctg gggcctgccc ctcactcatg gggaaccatc cctcacccgt 4380
gctgaatttg tttgagagca gcaaatccta cttttagtac agatgtgaga atttgaggca 4440
ttagtccaac aagtttttca gcctagaatt tgttttcctt tcccactacc catcaaggga 4500
tctggttact cagctagttc ccatcagctc tggctgtggt ctcggctgag tggccttggg 4560
gttatgtaag aggtagtggg aggggaagag agctgagggc tgcagcatat ggtccacctg 4620

ggcttggcct ttgggaatag gcagccctgg ctctctctga atccttagaa attacacggc 4680
tatttgatcc tggaaagatc gtgcagagca cacctgagtg tcatacagcc tggtctgagg 4740
tagtggggtt ggagatgagg tgggttgggg cacagtggtg tttagctcag gtaccaggtg 4800
gggaggttta gactttctgc tttaaaagga atgattagag cctggtctgg cgtttctttt 4860
gctggtccaa cacaccttga ccactttcat ccaggttttg ccaggtcctt gggtgagatc 4920
tgggctctct tccaggctgc acagacattt tcagaggtcc cctctgtgtg tgcaaaccta 4980
ggcaagccag gtgcctccct gtgaaacagg agaatgttgt gtagtcagag agtgacagga 5040
cctcctgagg gatttggggg aggatgggga tttgacagaa agaggccaga actggagaga 5100
gacagaacca gtctacgttg cagctctgtc ccccttagta gctatttgag tgtgaggaag 5160
ttactgaact tctgtttccc acatgagaaa tggtgataat agattcagcc ttgcagagta 5220
gtcgagtggg ttttctaagc ttatgttgta atttctcttg ggtacagagc acccagcacc 5280
gtgtagaatc ttcataagtg ttagctgtta ctgtggtaca acattactta aaggaagttg 5340
gaagagttaa ctccgcaaat ctggggaccc taagaggctg tgtgatgcct cagcacttga 5400
gcccacgtgg aaaggctgtg gccagggccc tgacctgctg tgtctgcagt ggggttgtcc 5460
cagccctcat gggcagctga ccttgagttc tggccttatt ttccccccctc tcagggcttg 5520
gagaagatct ctgcctacat gaagtccagc cgcttcctcc caaaacctct gtacacaagg 5580
gtggctgtct ggggcaacaa gtaatgcctt gaaggccagg aggtgggagt gaggagccca 5640
tactcagcct gctgcccagg ctgtgcagcg cagctggact ctgcatccca gcacctgcct 5700
cctcgttcct ttctcctgtt tattcccatc tttaccccca agactttatt gggcctcttc 5760
acttccccta aacccctgtc ccatgcaggc cctttgaagc ctcagctacc cactttcctt 5820
catgaacatc cccctcccaa cactaccctt cctgcacta aagccagcct gaccttcctt 5880
cctgttagtg gttgtatctg ctttgaaggg cctacctggc ccctcgcctg tggagctcag 5940
ccctgagctg tccccgtgtt gcatgacagc attgactggt ttacaggccc tgctcctgca 6000
gcatggcccc tgccttaggc ctacctgatc aaaataaagc ctcagccaca tttgctatag 6060
tcttgtctta tttgctcctg gc 6082

<210> 19

<211> 5102

<212> DNA

<213> Homo sapiens

<400> 19
gcttctgcga ctccagttgt gagagccgca agggcatggg aattgacgcc actcaccgac 60
ccccagtctc aatctcaacg ctgtgaggaa acctcgactt tgccaggtcc ccaagggcag 120
cggggctcgg cgagcgaggc acccttctcc gtccccatcc caatccaagc gctcctggca 180
ctgacgacgc caagagactc gagtgggagt taaagcttcc agtgagggca gcaggtgtcc 240
aggccgggcc tgcgggttcc tgttgacgtc ttgccctagg caaaggtccc agttccttct 300
cggagccggc tgtcccgcgc cactggaaac cgcacctccc cgcagcatgg gcaccagcct 360
cagcccgaac gacccttggc cgctaaaccc gctgtccatc cagcagacca cgctcctgct 420
actcctgtcg gtgctggcca ctgtgcatgt gggccagcgg ctgctgaggc aacggaggcg 480
gcagctccgg tccgcgcccc cgggcccgtt tgcgtggcca ctgatcggaa acgcggcggc 540
ggtgggccag gcggctcacc tctcgttcgc tcgcctggcg cggcgctacg gcgacgtttt 600
ccagatccgc ctgggcagct gccccatagt ggtgctgaat ggcgagcgcg ccatccacca 660
ggccctggtg cagcagggct cggccttcgc cgaccggccg gccttcgcct ccttccgtgt 720
ggtgtccggc ggccgcagca tggctttcgg ccactactcg gagcactgga aggtgcagcg 780
gcgcgcagcc cacagcatga tgcgcaactt cttcacgcgc cagccgcgca gccgccaagt 840
cctcgagggc cacgtgctga gcgaggcgcg cgagctggtg gcgctgctgg tgcgcggcag 900
cgcggacggc gccttcctcg acccgaggcc gctgaccgtc gtggccgtgg ccaacgtcat 960
gagtgccgtg tgtttcggct gccgctacag ccacgacgac cccgagttcc gtgagctgct 1020
cagccacaac gaagagttcg ggcgcacggt gggcgcgggc agcctggtgg acgtgatgcc 1080
ctggctgcag tacttcccca acccggtgcg caccgttttc cgcgaattcg agcagctcaa 1140
ccgcaacttc agcaacttca tcctggacaa gttcttgagg cactgcgaaa gccttcggcc 1200
cggggccgcc ccccgcgaca tgatggacgc ctttatcctc tctgcggaaa agaaggcggc 1260
cggggactcg cacggtggtg gcgcgcggct ggatttggag aacgtaccgg ccactatcac 1320
tgacatcttc ggcgccagcc aggacaccct gtccaccgcg ctgcagtggc tgctcctcct 1380
cttcaccagg tatcctgatg tgcagactcg agtgcaggca gaattggatc aggtcgtggg 1440
gagggaccgt ctgccttgta tgggtgacca gcccaacctg ccctatgtcc tggccttcct 1500
ttatgaagcc atgcgcttct ccagctttgt gcctgtcact attcctcatg ccaccactgc 1560
caacacctct gtcttgggct accacattcc caaggacact gtggtttttg tcaaccagtg 1620

gtctgtgaat catgacccag tgaagtggcc taacccggag aactttgatc cagctcgatt 1680
cttggacaag gatggcctca tcaacaagga cctgaccagc agagtgatga ttttttcagt 1740
gggcaaaagg cggtgcattg gcgaagaact ttctaagatg cagcttttc tcttcatctc 1800
catcctggct caccagtgcg atttcagggc caacccaaat gagcctgcga aaatgaattt 1860
cagttatggt ctaaccatta aacccaagtc atttaaagtc aatgtcactc tcagagagtc 1920
catggagctc cttgatagtg ctgtccaaaa tttacaagcc aaggaaactt gccaataaga 1980
agcaagaggc aagctgaaat tttagaaata ttcacatctt cggagatgag gagtaaaatt 2040
cagtttttt ccagttcctc ttttgtgctg cttctcaatt agcgtttaag gtgagcataa 2100
atcaactgtc catcaggtga ggtgtgctcc atacccagcg gttcttcatg agtagtgggc 2160
tatgcaggag cttctgggag attttttga gtcaaagact taaagggccc aatgaattat 2220
tatatacata ctgcatcttg gttatttctg aaggtagcat tctttggagt taaaatgcac 2280
atatagacac atacacccaa acacttacac caaactactg aatgaagaag tattttggta 2340
accaggccat ttttggtggg aatccaagat tggtctccca tatgcagaaa tagacaaaaa 2400
gtatattaaa caaagtttca gagtatattg ttgaagagac agagacaagt aatttcagtg 2460
taaagtgtgt gattgaaggt gataagggaa aagataaaga ccagaaattc ccttttcacc 2520
ttttcaggaa aataacttag actctagtat ttatgggtgg atttatcctt ttgccttctg 2580
gtatacttcc ttacttttaa ggataaatca taaagtcagt tgctcaaaaa gaaatcaata 2640
gttgaattag tgagtatagt ggggttccat gagttatcat gaattttaaa gtatgcatta 2700
ttaaattgta aaactccaag gtgatgttgt acctcttttg cttgccaaag tacagaattt 2760
gaattatcag caaagaaaaa aaaaaaagcc agccaagctt taaattatgt gaccataatg 2820
tactgatttc agtaagtctc ataggttaaa aaaaaaagtc accaaatagt gtgaaatata 2880
ttacttaact gtccgtaagc agtatattag tattatcttg ttcaggaaaa ggttgaataa 2940
tatatgcctt gtgtaatatt gaaaattgaa aagtacaact aacgcaacca agtgtgctaa 3000
aaatgagctt gattaaatca accacctatt tttgacatgg aaatgaagca gggtttcttt 3060
tcttcactca aattttggcg aatctcaaaa ttagatccta agatgtgttc ttatttttat 3120
aacatcttta ttgaaattct atttataata cagaatcttg ttttgaaaat aacctaatta 3180
atatattaaa attccaaatt catggcatgc ttaaatttta actaaatttt aaagccattc 3240
tgattattga gttccagttg aagttagtgg aaatctgaac attctcctgt ggaaggcaga 3300
gaaatctaag ctgtgtctgc ccaatgaata atggaaaatg ccatgaatta cctggatgtt 3360
ctttttacga ggtgacaaga gttggggaca gaactcccat tacaactgac caagtttctc 3420

ttctagatga ttttttgaaa gttaacatta atgcctgctt tttggaaagt cagaatcaga 3480
agatagtctt ggaagctgtt tggaaaagac agtggagatg aggtcagttg tgtttttaa 3540
gatggcaatt actttggtag ctgggaaagc ataaagctca aatgaaatgt atgcattcac 3600
atttagaaaa gtgaattgaa gtttcaagtt ttaaagttca ttgcaattaa acttccaaag 3660
aaagttctac agtgtcctaa gtgctaagtg cttattacat tttattaagc tttttggaat 3720
ctttgtacca aaattttaaa aaagggagtt tttgatagtt gtgtgtatgt gtgtgtgggg 3780
tgggggggatg gtaagagaaa agagagaaac actgaaaaga aggaaagatg gttaaacatt 3840
ttcccactca ttctgaatta attaatttgg agcacaaaat tcaaagcatg gacatttaga 3900
agaaagatgt ttggcgtagc agagttaaat ctcaaatagg ctattaaaaa agtctacaac 3960
atagcagatc tgttttgtgg tttggaatat taaaaaactt catgtaattt tattttaaaa 4020
tttcatagct gtacttcttg aatataaaaa atcatgccag tatttttaaa ggcattagag 4080
tcaactacac aaagcaggct tgcccagtac atttaaattt tttggcactt gccattccaa 4140
aatattatgc cccaccaagg ctgagacagt gaatttgggc tgctgtagcc tatttttta 4200
gattgagaaa tgtgtagctg caaaaataat catgaaccaa tctggatgcc tcattatgtc 4260
aaccaggtcc agatgtgcta taatctgttt ttacgtatgt aggcccagtc gtcatcagat 4320
gcttgcggca aaagaaagct gtgtttatat ggaagaaagt aaggtgcttg gagtttacct 4380
ggcttattta atatgcttat aacctagtta aagaaaggaa aagaaaacaa aaaacgaatg 4440
aaaataactg aatttggagg ctggagtaat cagattactg ctttaatcag aaaccctcat 4500
tgtgtttcta ccggagagag aatgtatttg ctgacaacca ttaaagtcag aagttttact 4560
ccaggttatt gcaataaagt ataatgttta ttaaatgctt catttgtatg tcaaagcttt 4620
gactctataa gcaaattgct tttttccaaa acaaaaagat gtctcaggtt tgttttgtga 4680
attttctaaa agctttcatg tcccagaact tagcctttac ctgtgaagtg ttactacagc 4740
cttaatattt tcctagtaga tctatattag atcaaatagt tgcatagcag tatatgttaa 4800
tttgtgtgtt tttagctgtg acacaactgt gtgattaaaa ggtatacttt agtagacatt 4860
tataactcaa ggataccttc ttatttaatc ttttcttatt tttgtacttt atcatgaatg 4920
cttttagtgt gtgcataata gctacagtgc atagttgtag acaaagtaca ttctggggaa 4980
acaacattta tatgtagcct ttactgtttg atataccaaa ttaaaaaaaa attgtatctc 5040
attacttata ctgggacacc attaccaaaa taataaaaat cactttcata atcttgaaaa 5100
aa 5102

<210> 20

<211> 1524

<212> DNA

<213> Homo sapiens

<400> 20
gcagagcaca gcatcgtcgg gaccagactc gtctcaggcc agttgcagcc ttctcagcca 60

aacgccgacc aaggaaaact cactaccatg agaattgcag tgatttgctt ttgcctccta 120

ggcatcacct gtgccatacc agttaaacag gctgattctg gaagttctga ggaaaagcag 180

ctttacaaca aatacccaga tgctgtggcc acatggctaa accctgaccc atctcagaag 240

cagaatctcc tagccccaca gacccttcca agtaagtcca acgaaagcca tgaccacatg 300

gatgatatgg atgatgaaga tgatgatgac catgtggaca gccaggactc cattgactcg 360

aacgactctg atgatgtaga tgacactgat gattctcacc agtctgatga gtctcaccat 420

tctgatgaat ctgatgaact ggtcactgat tttcccacgg acctgccagc aaccgaagtt 480

ttcactccag ttgtccccac agtagacaca tatgatggcc gaggtgatag tgtggtttat 540

ggactgaggt caaaatctaa gaagtttcgc agacctgaca tccagtaccc tgatgctaca 600

gacgaggaca tcacctcaca catggaaagc gaggagttga atggtgcata caaggccatc 660

cccgttgccc aggacctgaa cgcgccttct gattgggaca gccgtgggaa ggacagttat 720

gaaacgagtc agctggatga ccagagtgct gaaacccaca gccacaagca gtccagatta 780

tataagcgga aagccaatga tgagagcaat gagcattccg atgtgattga tagtcaggaa 840

ctttccaaag tcagccgtga attccacagc catgaatttc acagccatga agatatgctg 900

gttgtagacc ccaaaagtaa ggaagaagat aaacacctga aatttcgtat ttctcatgaa 960

ttagatagtg catcttctga ggtcaattaa aaggagaaaa aatacaattt ctcactttgc 1020

atttagtcaa aagaaaaaat gctttatagc aaaatgaaag agaacatgaa atgcttcttt 1080

ctcagtttat tggttgaatg tgtatctatt tgagtctgga aataactaat gtgtttgata 1140

attagtttag tttgtggctt catggaaact ccctgtaaac taaaagcttc agggttatgt 1200

ctatgttcat tctatagaag aaatgcaaac tatcactgta ttttaatatt tgttattctc 1260

tcatgaatag aaatttatgt agaagcaaac aaaatacttt tacccactta aaaagagaat 1320

ataacatttt atgtcactat aatcttttgt tttttaagtt agtgtatatt ttgttgtgat 1380

tatctttttg tggtgtgaat aaatctttta tcttgaatgt aataagaatt tggtggtgtc 1440

aattgcttat ttgttttccc acggttgtcc agcaattaat aaaacataac ctttttttact 1500

gcctaaaaaa aaaaaaaaaa aaaa 1524

<210> 21

<211> 478

<212> DNA

<213> Homo sapiens

<400> 21
cttcaagacc attgtggcca aggagatctg tgctgacccc aagcagaagt gggttcagga 60

ttccatggac cacctggaca agcaaaccca aactccgaag acttgaacac ccactccaca 120

acccaagaat ctgcagctaa cttattttcc cctagctttc cccagacacc ttgttttatt 180

ttattataat gaatttgttt gttgatgtga aacattatgc cttaagtaat gttaattctt 240

atttaagtta ttgatgtttt aagtatcttt catggtacta gtgtttttta gatacagaga 300

cttggggaaa ttgcttttcc tcttttgaac cacagttcta cccctgggat gttttgaggg 360

tctttgcaag aatcattaat acaaagaatt ttttttaaca ttccaatgca ttgctaaaat 420

attattgtgg aaatgaatat tttgtaacta ttacaccaaa taaatatatt tttgtaca 478

<210> 22

<211> 1778

<212> DNA

<213> Homo sapiens

<400> 22
tagaagttta caatgaagtt tcttctaata ctgctcctgc aggccactgc ttctggagct 60

cttcccctga acagctctac aagcctggaa aaaaataatg tgctatttgg tgagagatac 120

ttagaaaaat tttatggcct tgagataaac aaacttccag tgacaaaaat gaaatatagt 180

ggaaacttaa tgaaggaaaa aatccaagaa atgcagcact tcttgggtct gaaagtgacc 240

gggcaactgg acacatctac cctggagatg atgcacgcac ctcgatgtgg agtccccgat 300

ctccatcatt tcagggaaat gccaggggggg cccgtatgga ggaaacatta tatcacctac 360

agaatcaata attacacacc tgacatgaac cgtgaggatg ttgactacgc aatccggaaa 420

gctttccaag tatggagtaa tgttaccccc ttgaaattca gcaagattaa cacaggcatg 480
gctgacattt tggtggtttt tgcccgtgga gctcatggag acttccatgc ttttgatggc 540
aaaggtggaa tcctagccca tgcttttgga cctggatctg gcattggagg ggatgcacat 600
ttcgatgagg acgaattctg gactacacat tcaggaggca caaacttgtt cctcactgct 660
gttcacgaga ttggccattc cttaggtctt ggccattcta gtgatccaaa ggctgtaatg 720
ttccccacct acaaatatgt cgacatcaac acatttcgcc tctctgctga tgacatacgt 780
ggcattcagt ccctgtatgg agacccaaaa gagaaccaac gcttgccaaa tcctgacaat 840
tcagaaccag ctctctgtga ccccaatttg agttttgatg ctgtcactac cgtgggaaat 900
aagatctttt tcttcaaaga caggttcttc tggctgaagg tttctgagag accaaagacc 960
agtgttaatt taatttcttc cttatggcca accttgccat ctggcattga agctgcttat 1020
gaaattgaag ccagaaatca agtttttctt tttaaagatg acaaatactg gttaattagc 1080
aatttaagac cagagccaaa ttatcccaag agcatacatt cttttggttt tcctaacttt 1140
gtgaaaaaaa ttgatgcagc tgtttttaac ccacgttttt ataggaccta cttctttgta 1200
gataaccagt attggaggta tgatgaaagg agacagatga tggaccctgg ttatcccaaa 1260
ctgattacca agaacttcca aggaatcggg cctaaaattg atgcagtctt ctattctaaa 1320
aacaaatact actatttctt ccaaggatct aaccaatttg aatatgactt cctactccaa 1380
cgtatcacca aaacactgaa aagcaatagc tggtttggtt gttagaaatg gtgtaattaa 1440
tggtttttgt tagttcactt cagcttaata agtatttatt gcatatttgc tatgtcctca 1500
gtgtaccact acttagagat atgtatcata aaaataaaat ctgtaaacca taggtaatga 1560
ttatataaaa tacataatat ttttcaattt tgaaaactct aattgtccat tcttgcttga 1620
ctctactatt aagtttgaaa atagttacct tcaaagcaag ataattctat ttgaagcatg 1680
ctctgtaagt tgcttcctaa catccttgga ctgagaaatt atacttactt ctggcataac 1740
taaaattaag tatatatatt ttggctcaaa taaaattg 1778

<210> 23

<211> 599

<212> DNA

<213> Homo sapiens

<400> 23
ttttttttttt tgctgttacg attatttatt gttactattc actgctttat tcccatggac 60

ccatgggtta gagctccata cataaagtag gtgctcatta aatgttggct ggactgagtc 120
agcccctcca tcttggccct tgagtcctca gaaaagcagt gccatccatg agaagcaact 180
tgaattgtat tttttattga aaagaattca ggctagagtt gggaggagga tgcaagagct 240
actgggaagg gggagctcag tctgaacctg ggggatcagg ggagtagggg actctcccct 300
tgtccactga tgggggggtct ggctgttact cctctccctt cagcacagaa agaacttggt 360
cagtaaaaat gcctgtgtaa gtgctcatgg ctgctgtgct tttgctgtac aagtccctga 420
gtttctcatc tacagcgggc aggtatgtct tctcgtacag gttctgggcg gctgtctttg 480
ctgactccca gtaactggag agagattcct tcacctgggt gaggaaggtc gggctaggca 540
tctcatcttg ctggggctgt tgggtccctg gactcaaatc ccaataccag gaggacaag 599

<210> 24
<211> 1270
<212> DNA
<213> Homo sapiens

<400> 24
gcggacgtgg gcaggagggc tggaaaagcc ggcgctggag cggaacggga gtagctgcct 60
gggcgccaaa ggccgcggca ctcccacgtg gaccccgaag tcccgaaccc ggggatgggc 120
ccgcggctgc gaggggatct tctctggatc aagcaatggt ggtgaaaaat gtttcgcaag 180
ggcaaaaaac gacacagtag tagcagttcc caaagtagcg aaatcagtac taagagcaag 240
tctgtggatt ctagccttgg gggtctttca cgatccagca ctgtggccag cctcgacaca 300
gattccacca aaagctcagg acaaagcaac aataattcag atacctgtgc agaatttcga 360
ataaaatatg ttggtgccat tgagaaactg aaactctccg agggaaaagg ccttgaaggg 420
ccattagacc tgataaatta tatagacgtt gcccagcaag atggaaagtt gccttttgtt 480
cctccggagg aagaatttat tatgggagtt tccaagtatg gcataaaagt atcaacatca 540
gatcaatatg aacaagcaca agccatttgc aaggttttat ccaccgcttt tgactctgta 600
ttaacatctg agaaaccctg aatcctgcaa tcaagtagaa gtcaacttca tctgaaagtt 660
cagctgtttt caaactgcaa tgctgaaatg ttatgcaaat aatgaagtta tcccttgctc 720
tagattttct gaagaaaatg gattgtgtaa aatgctgatc atttgtttat taaaatgtgt 780
cctattacac agtgagttaa ctctcaatga agtcatctat tttctgggct aaaaaacttc 840

atttgtcttt ttcaacttct aataagctta acctaagtgt cacgaagacg agatgtcaca 900
gaggtccact cagtgacaca cactgaaggc ctgagggaag actgaggaca tgggctcagt 960
ggtggcttcc cagtcatggt atcactggca tggacctctg tccggcagag gtgtggactg 1020
gagaccagga ttcatgctgg tctggaacaa tgacattgcc aacttaagac acacaaagca 1080
gattttcaga agtgtctggt caagataaca tgctggccaa ccacaattcc tagagttaag 1140
agaaccttaa aagattaccg ctcatgctaa aagtatgtaa agatcccatg tacagtatga 1200
tagtgtactt tttttaaagg actgtcaata tacaaaactt taaagattaa aaacattaaa 1260
aataaaaaaa 1270

<210> 25
<211> 2390
<212> DNA
<213> Homo sapiens

<400> 25
ccagcctcgg agggagggat tagaagccgc tagacttttt ttcctcccct ctcagtagca 60
cggagtccga attaattgga tttcattcac tggggaggaa caaaaactat ctgggcagct 120
tcattgagag agattcattg acactaagag ccagcgctgc agctggtgca gagagaacct 180
ccggctttga cttctgtctc gtctgcccca aggccgctag cctcggcttg ggaaggcgag 240
gcggaattaa accccgctcc gagagcgcac gttcgcgcgc ggtgcgtcgg ccattgcctg 300
ccccgagggg cgtctggtag gcaccccgcc ctctcccgca gctcgacccc catgatagat 360
acgctcagac ccgtgccctt cgcgtcggaa atggcgatca gcaagacggt ggcgtggctc 420
aacgagcagc tggagctggg caacgagcgg ctgctgctga tggactgccg gccgcaggag 480
ctatacgagt cgtcgcacat cgagtcggcc atcaacgtgg ccatcccggg catcatgctg 540
cggcgcctgc agaagggtaa cctgccggtg cgcgcgctct tcacgcgcgg cgaggaccgg 600
gaccgcttca cccggcgctg tggcaccgac acagtggtgc tctacgacga gagcagcagc 660
gactggaacg agaatacggg cggcgagtcg ttgctcgggc tgctgctcaa gaagctcaag 720
gacgagggct gccgggcgtt ctacctggaa ggtggcttca gtaagttcca agccgagttc 780
tccctgcatt gcgagaccaa tctagacggc tcgtgtagca gcagctcgcc gccgttgcca 840
gtgctggggc tcgggggcct gcggatcagc tctgactctt cctcggacat cgagtctgac 900
cttgaccgag accccaatag tgcaacagac tcggatggta gtccgctgtc caacagccag 960

ccttccttcc cagtggagat cttgcccttc ctctacttgg gctgtgccaa agactccacc 1020
aacttggacg tgttggagga attcggcatc aagtacatct tgaacgtcac ccccaatttg 1080
ccgaatctct ttgagaacgc aggagagttt aaatacaagc aaatccccat ctcggatcac 1140
tggagccaaa acctgtccca gtttttccct gaggccattt ctttcataga tgaagcccgg 1200
ggcaagaact gtggtgtctt ggtacattgc ttggctggca ttagccgctc agtcactgtg 1260
actgtggctt accttatgca gaagctcaat ctgtcgatga acgatgccta tgacattgtc 1320
aaaatgaaaa aatccaacat atcccctaac ttcaacttca tgggtcagct gctggacttc 1380
gagaggacgc tgggactcag cagcccatgt gacaacaggg ttccagcaca gcagctgtat 1440
tttaccaccc cttccaacca gaatgtatac caggtggact ctctgcaatc tacgtgaaag 1500
accccacacc cctccttgct ggaatgtgtc tggcccttca gcagtttctc ttggcagcat 1560
cagctgggct gctttctttg tgtgtggccc caggtgtcaa aatgacacca gctgtctgta 1620
ctagacaagg ttaccaagtg cggaattggt taatactaac agagagattt gctccattct 1680
ctttggaata acaggacatg ctgtatagat acaggcagta ggtttgctct gtacccatgt 1740
gtacagccta cccatgcagg gactgggatt cgaggacttc caggcgcata gggtagaacc 1800
aaatgatagg gtaggagcat gtgttcttta gggccttgta aggctgtttc cttttgcatc 1860
tggaactgac tatataattg tcttcaagtg aagactaatt caattttgca tatagaggag 1920
ccaaagagag atttcagctc tgtatttgtg gtatcagtgt tggaaaagag aaatctgata 1980
ctccatttgg attattgtaa atatttgatc ttgaatcact tgacagtgtt tgtttgaatt 2040
gtgtttgttt tttcctttga tgggcttaaa agaaattatc caaagggaga aagagcagta 2100
tgccacttct taaaacagaa caaaacaaaa aaagaaaatt gtgctctgtt gtaatccaaa 2160
ggggagattt gcagcatgct tgactttacc aattctgatg acatctttac ggacactatt 2220
atcactaaga ccttgttatg gcgaagtctt tagtcttttt catgtatttt cctcatgatt 2280
ttttctcttt atgtagtttg agtatgcctt acctttgtaa atatttttgc ttgtgttgtc 2340
gcaaagggga taatctggga aagacaccaa atcatgggct cactttaaaa 2390

<210> 26

<211> 3871

<212> DNA

<213> Homo sapiens

<400> 26
gcggattgtt gctccgttgt acctgctggg gaattcacct cgttactgct tgatatcttc 60
cacccttac aaaatcagaa aagttgtgtt ttctaatacc aaagaggagg tttggctttc 120
tgtgggtgat tcccagacac tgaagtgcaa agaagagacc ctcctagaaa agtaaaatat 180
gactaaaagc aatggagaag agcccaagat ggggggcagg atggagagat tccagcaggg 240
agtccgtaaa cgcacacttt tggccaagaa gaaagtgcag aacattacaa aggaggatgt 300
taaaagttac ctgtttcgga atgcttttgt gctgctcaca gtcaccgctg tcattgtggg 360
tacaatcctt ggatttaccc tccgaccata cagaatgagc taccgggaag tcaagtactt 420
ctcctttcct ggggaacttc tgatgaggat gttacagatg ctggtcttac cacttatcat 480
ctccagtctt gtcacaggaa tggcggcgct agatagtaag gcatcaggga agatgggaat 540
gcgagctgta gtctattata tgactaccac catcattgct gtggtgattg gcataatcat 600
tgtcatcatc atccatcctg ggaagggcac aaaggaaaac atgcacagag aaggcaaaat 660
tgtacgagtg acagctgcag atgccttcct ggacttgatc aggaacatgt tccctccaaa 720
tctggtagaa gcctgcttta aacagtttaa aaccaactat gagaagagaa gctttaaagt 780
gcccatccag gccaacgaaa cgcttgtggg tgctgtgata aacaatgtgt ctgaggccat 840
ggagactctt acccgaatca cagaggagct ggtcccagtt ccaggatctg tgaatggagt 900
caatgccctg ggtctagttg tcttctccat gtgcttcggt tttgtgattg gaaacatgaa 960
ggaacagggg caggccctga gagagttctt tgattctctt aacgaagcca tcatgagact 1020
ggtagcagta ataatgtggt atgccccgt gggtattctc ttcctgattg ctgggaagat 1080
tgtggagatg gaagacatgg gtgtgattgg ggggcagctt gccatgtaca ccgtgactgt 1140
cattgttggc ttactcattc acgcagtcat cgtcttgcca ctcctctact tcttggtaac 1200
acggaaaaac ccttgggttt ttattggagg gttgctgcaa gcactcatca ccgctctggg 1260
gacctcttca agttctgcca ccctacccat caccttcaag tgcctggaag agaacaatgg 1320
cgtggacaag cgcgtcacca gattcgtgct ccccgtagga gccaccatta acatggatgg 1380
gactgccctc tatgaggctt tggctgccat tttcattgct caagttaaca actttgaact 1440
gaacttcgga caaattatta caatcagcat cacagccaca gctgccagta ttggggcagc 1500
tggaattcct caggcgggcc tggtcactat ggtcattgtg ctgacatctg tcggcctgcc 1560
cactgacgac atcacgctca tcatcgcggt ggactggttc ctggatcgcc tccggaccac 1620
caccaacgta ctgggagact ccctgggagc tgggattgtg gagcacttgt cacgacatga 1680

actgaagaac agagatgttg aaatgggtaa ctcagtgatt gaagagaatg aaatgaagaa 1740
accatatcaa ctgattgcac aggacaatga aactgagaaa cccatcgaca gtgaaaccaa 1800
gatgtagact aacataaaga aacactttct tgagcaccag gtgttaaaaa ccattataaa 1860
atctttccat ctcattacag ctcattcgct ccagcaagcc cgtcatcttc cctttcctcc 1920
cttctgataa gactggaaaa tagtcctcca aaacacaagg gaggattttg ggtggccaaa 1980
gtgtacaatt ttcatcccac aattgaaatt tttaaatcat ttcatgttag tcttaccgaa 2040
taaggtacca agatcacaaa tagtgttgat cagatcttac aagtttatgt ggcacacaat 2100
tcctataaat gtgatttttt tatataagtt aaagagacaa atagtaggct aaaaacattt 2160
taaaatcaac ttttgaaatt taaaaatctt tcagaataca attcagtttt agtttcaaaa 2220
tgttaacaac ttgaattaca accggttatc agttggacag taagatttta tccctttctc 2280
ttctgactgg tatacctatt tcattagtag ctaggtgcac atatacatct agcacagctg 2340
tgaggacaga cagaaggcaa agtttccatg tggccttgag caagtcccat ctcacctcta 2400
ggcctcagtg tcctcatcta taaaatgagg gacttcccta gaagtcttca tggtctcttc 2460
cagcccagac atcctgtgat gtcatgaaag cacctgccct ctgtttcccc tcagaacacc 2520
ctgtaccatc catggagcac gaggccttca gaaaagacac ttcaatggga gtgaacattt 2580
ctaactaagg acaggatggc tgtgtgtggt ggtcaccagg tcctgtgagc aaagtgcagg 2640
ttatgcaagt cgccaggcag gaggccattc caggagtggg attattcatc aaactctttg 2700
cccagttcat cccaatgggg gaagtattcc cttctttcct actctgggaa gaatgtctcc 2760
tgccactcct caactgatga tagacttcga aaacagatga gaagactagc agctagcaag 2820
ggtgcttgta gtcacactgt ggaacactaa agagctagga aagagttgag cacaggcaac 2880
attacaaaca aaggatttga aaacaccaag agtacaggtc ttctttaagg aagaataaaa 2940
aagaagaggt tcatttttct ggctttttt ttcacctgaa acacttttc tcgagtccaa 3000
aatcattccc cccgtgaagt ctgcttacca aaacataaga cgacttatat atttgaaaga 3060
agtcaaatga atgagctctc taatagaagt ccatgagttg agtgggtatt tcttatttga 3120
aagtgttttt ctttaatcaa aagtccttag aatgagggaa acaaaatatt tatttgtttt 3180
ggaatcccac ttatcaaatc attcaaaact ttcagctgga gtggggtttg cttttgtttt 3240
gtttgtgtcc ataagagaaa tggtagaaga tgaatcagta tgaagacact gtcaatgagg 3300
ttatgagaaa aaaacagcag gggcattagt ttcaggcaag gcagctccca ggtttagaga 3360
ttaatttta ccccctaagg aatatccagt caaagacgct gagtgggagc tgtcaggcag 3420
tagcagctgt gtttgagttt ctggctgaaa atggtgaaga atggacttaa ttatgctaac 3480

aaactgaaaa atctagacat agatcctctg atatacaatt agagatattt ttatatagac 3540
cccaagcatt ctgtgcataa aagttaacat taggctgtgg tgcagtaacc atttaatgtc 3600
gaggctctat ttcggaaata cactacaaat gttaaagtac gtggctgtcc tcttaagaca 3660
ctagtagagc aaagacttaa tcatatcaac ttaattctgt tacacaatat gtgtttttta 3720
atatactaac catttcttat ggaaaggtcc tgtggggagc ccatcctctc gccaagccat 3780
cacaggctct gcatacacat gcactcagtg tggactggga agcattactt tgtagatgta 3840
ttttcaataa agaaaaaaat agttttacat t 3871

<210> 27
<211> 1433
<212> DNA
<213> Homo sapiens

<400> 27
attcggggcg agggaggagg aagaagcgga ggaggcggct cccgctcgca gggccgtgca 60
cctgcccgcc cgcccgctcg ctcgctcgcc cgccgcgccg cgctgccgac cgccagcatg 120
ctgccgagag tgggctgccc cgcgctgccg ctgccgccgc cgccgctgct gccgctgctg 180
ccgctgctgc tgctgctact gggcgcgagt ggcggcggcg gcggggcgcg cgcggaggtg 240
ctgttccgct gcccgccctg cacacccgag cgcctggccg cctgcgggcc cccgccggtt 300
gcgccgcccg ccgcggtggc cgcagtggcc ggaggcgccc gcatgccatg cgcggagctc 360
gtccgggagc cgggctgcgg ctgctgctcg gtgtgcgccc ggctggaggg cgaggcgtgc 420
ggcgtctaca ccccgcgctg cggccagggg ctgcgctgct atccccaccc gggctccgag 480
ctgcccctgc aggcgctggt catgggcgag ggcacttgtg agaagcgccg ggacgccgag 540
tatggcgcca gcccggagca ggttgcagac aatggcgatg accactcaga aggaggcctg 600
gtggagaacc acgtggacag caccatgaac atgttgggcg ggggaggcag tgctggccgg 660
aagcccctca agtcgggtat gaaggagctg gccgtgttcc gggagaaggt cactgagcag 720
caccggcaga tgggcaaggg tggcaagcat caccttggcc tggaggagcc caagaagctg 780
cgaccacccc ctgccaggac tccctgccaa caggaactgg accaggtcct ggagcggatc 840
tccaccatgc gccttccgga tgagcggggc cctctggagc acctctactc cctgcacatc 900
cccaactgtg acaagcatgg cctgtacaac ctcaaacagt gcaagatgtc tctgaacggg 960

cagcgtgggg agtgctggtg tgtgaacccc aacaccggga agctgatcca gggagccccc 1020

accatccggg gggaccccga gtgtcatctc ttctacaatg agcagcagga ggcttgcggg 1080

gtgcacaccc agcggatgca gtagaccgca gccagccggt gcctggcgcc cctgcccccc 1140

gcccctctcc aaacaccggc agaaaacgga gagtgcttgg gtggtgggtg ctggaggatt 1200

ttccagttct gacacacgta tttatatttg gaaagagacc agcaccgagc tcggcacctc 1260

cccggcctct ctcttcccag ctgcagatgc cacacctgct ccttcttgct ttccccgggg 1320

gaggaagggg gttgtggtcg gggagctggg gtacaggttt ggggaggggg aagagaaatt 1380

tttatttttg aacccctgtg tccctttttgc ataagattaa aggaaggaaa agt 1433

<210> 28

<211> 1011

<212> DNA

<213> Homo sapiens

<400> 28
ggtttatttt ccagatgcaa tcaatgcccc agtcacctgc tgttataact tcaccaatag 60

gaagatctca gtgcagaggc tcgcgagcta tagaagaatc accagcagca agtgtcccaa 120

acaagctgtg atgtgagttc agcacaccaa ccttccctgg cctgaagttc ttccttgtgg 180

agcaagggac aagcctcata aacctagagt cagagagtgc actatttaac ttaatgtaca 240

aaggttccca atgggaaaac tgaggcacca agggaaaaag tgaaccccaa catcactctc 300

cacctgggtg cctattcaga acacccaatt tctttagctt gaagtcagga tggctccacc 360

tggacaccta taggagcagt ttgccctggg ttccctcctt ccacctgcgt tcctcctcta 420

gctcccatgg cagccctttg gtgcagaatg ggctgcactt ctagaccaaa actgcaaagg 480

aacttcatct aactctgtcc tccctcccca cagcttacag accattgtgg caaggagatc 540

tgtgctgacc ccaagcagaa gtgggttcag gattccatgg accacctgga caagcaaacc 600

caaactccga agacttgaac actcactcca caacccaaga atctgcagct aacttatttt 660

tccctagctt tccccagaca ccttgtttat tttattataa tgaattttgt ttgttgatgt 720

gaaacattat gccttaagta atgttaattc ttatttaagt tattgatgtt ttaagtttat 780

ctttcatggt actagtgttt tttagataca gagacttggg gaaattgctt ttcctcttga 840

accacagttc taccctgggg atgttttgag ggtctttgca agaatcatta atacaaagaa 900

ttttttttaa cattccaatg cattgctaaa atattattgt ggaaatgaat attttgtaac 960

tattacacca aataaatata tttttgtaca aaaaaaaaaa aaaaaaaaaa a 1011

<210> 29

<211> 211

<212> PRT

<213> Homo sapiens

<400> 29

Met Gln Ser Ala Met Phe Leu Ala Val Gln His Asp Cys Arg Pro Met
1 5 10 15

Asp Lys Ser Ala Gly Ser Gly His Lys Ser Glu Glu Lys Arg Glu Lys
20 25 30

Met Lys Arg Thr Leu Leu Lys Asp Trp Lys Thr Arg Leu Ser Tyr Phe
35 40 45

Leu Gln Asn Ser Ser Thr Pro Gly Lys Pro Lys Thr Gly Lys Lys Ser
50 55 60

Lys Gln Gln Ala Phe Ile Lys Pro Ser Pro Glu Glu Ala Gln Leu Trp
65 70 75 80

Ser Glu Ala Phe Asp Glu Leu Leu Ala Ser Lys Tyr Gly Leu Ala Ala
85 90 95

Phe Arg Ala Phe Leu Lys Ser Glu Phe Cys Glu Glu Asn Ile Glu Phe
100 105 110

Trp Leu Ala Cys Glu Asp Phe Lys Lys Thr Lys Ser Pro Gln Lys Leu
115 120 125

Ser Ser Lys Ala Arg Lys Ile Tyr Thr Asp Phe Ile Glu Lys Glu Ala
130 135 140

Pro Lys Glu Ile Asn Ile Asp Phe Gln Thr Lys Thr Leu Ile Ala Gln
145 150 155 160

Asn Ile Gln Glu Ala Thr Ser Gly Cys Phe Thr Thr Ala Gln Lys Arg
165 170 175

Val Tyr Ser Leu Met Glu Asn Asn Ser Tyr Pro Arg Phe Leu Glu Ser
180 185 190

Glu Phe Tyr Gln Asp Leu Cys Lys Lys Pro Gln Ile Thr Thr Glu Pro
195 200 205

His Ala Thr
210

<210> 30

<211> 1170

<212> PRT

<213> Homo sapiens

<400> 30

Met Gly Leu Ala Trp Gly Leu Gly Val Leu Phe Leu Met His Val Cys
1 5 10 15

Gly Thr Asn Arg Ile Pro Glu Ser Gly Gly Asp Asn Ser Val Phe Asp
20 25 30

Ile Phe Glu Leu Thr Gly Ala Ala Arg Lys Gly Ser Gly Arg Arg Leu
35 40 45

Val Lys Gly Pro Asp Pro Ser Ser Pro Ala Phe Arg Ile Glu Asp Ala
50 55 60

Asn Leu Ile Pro Pro Val Pro Asp Asp Lys Phe Gln Asp Leu Val Asp
65 70 75 80

Ala Val Arg Ala Glu Lys Gly Phe Leu Leu Leu Ala Ser Leu Arg Gln
85 90 95

Met Lys Lys Thr Arg Gly Thr Leu Leu Ala Leu Glu Arg Lys Asp His
100 105 110

Ser Gly Gln Val Phe Ser Val Val Ser Asn Gly Lys Ala Gly Thr Leu
115 120 125

Asp Leu Ser Leu Thr Val Gln Gly Lys Gln His Val Val Ser Val Glu
130 135 140

```
Glu Ala Leu Leu Ala Thr Gly Gln Trp Lys Ser Ile Thr Leu Phe Val
145                 150                 155                 160

Gln Glu Asp Arg Ala Gln Leu Tyr Ile Asp Cys Glu Lys Met Glu Asn
                165                 170                 175

Ala Glu Leu Asp Val Pro Ile Gln Ser Val Phe Thr Arg Asp Leu Ala
            180                 185                 190

Ser Ile Ala Arg Leu Arg Ile Ala Lys Gly Gly Val Asn Asp Asn Phe
        195                 200                 205

Gln Gly Val Leu Gln Asn Val Arg Phe Val Phe Gly Thr Thr Pro Glu
    210                 215                 220

Asp Ile Leu Arg Asn Lys Gly Cys Ser Ser Ser Thr Ser Val Leu Leu
225                 230                 235                 240

Thr Leu Asp Asn Asn Val Val Asn Gly Ser Ser Pro Ala Ile Arg Thr
                245                 250                 255

Asn Tyr Ile Gly His Lys Thr Lys Asp Leu Gln Ala Ile Cys Gly Ile
            260                 265                 270

Ser Cys Asp Glu Leu Ser Ser Met Val Leu Glu Leu Arg Gly Leu Arg
        275                 280                 285

Thr Ile Val Thr Thr Leu Gln Asp Ser Ile Arg Lys Val Thr Glu Glu
    290                 295                 300

Asn Lys Glu Leu Ala Asn Glu Leu Arg Arg Pro Pro Leu Cys Tyr His
305                 310                 315                 320

Asn Gly Val Gln Tyr Arg Asn Asn Glu Glu Trp Thr Val Asp Ser Cys
                325                 330                 335

Thr Glu Cys His Cys Gln Asn Ser Val Thr Ile Cys Lys Lys Val Ser
            340                 345                 350

Cys Pro Ile Met Pro Cys Ser Asn Ala Thr Val Pro Asp Gly Glu Cys
        355                 360                 365
```

```
Cys Pro Arg Cys Trp Pro Ser Asp Ser Ala Asp Asp Gly Trp Ser Pro
    370                 375                 380

Trp Ser Glu Trp Thr Ser Cys Ser Thr Ser Cys Gly Asn Gly Ile Gln
385                 390                 395                 400

Gln Arg Gly Arg Ser Cys Asp Ser Leu Asn Asn Arg Cys Glu Gly Ser
                405                 410                 415

Ser Val Gln Thr Arg Thr Cys His Ile Gln Glu Cys Asp Lys Arg Phe
            420                 425                 430

Lys Gln Asp Gly Gly Trp Ser His Trp Ser Pro Trp Ser Ser Cys Ser
        435                 440                 445

Val Thr Cys Gly Asp Gly Val Ile Thr Arg Ile Arg Leu Cys Asn Ser
    450                 455                 460

Pro Ser Pro Gln Met Asn Gly Lys Pro Cys Glu Gly Glu Ala Arg Glu
465                 470                 475                 480

Thr Lys Ala Cys Lys Lys Asp Ala Cys Pro Ile Asn Gly Gly Trp Gly
                485                 490                 495

Pro Trp Ser Pro Trp Asp Ile Cys Ser Val Thr Cys Gly Gly Gly Val
            500                 505                 510

Gln Lys Arg Ser Arg Leu Cys Asn Asn Pro Ala Pro Gln Phe Gly Gly
        515                 520                 525

Lys Asp Cys Val Gly Asp Val Thr Glu Asn Gln Ile Cys Asn Lys Gln
    530                 535                 540

Asp Cys Pro Ile Asp Gly Cys Leu Ser Asn Pro Cys Phe Ala Gly Val
545                 550                 555                 560

Lys Cys Thr Ser Tyr Pro Asp Gly Ser Trp Lys Cys Gly Ala Cys Pro
                565                 570                 575

Pro Gly Tyr Ser Gly Asn Gly Ile Gln Cys Thr Asp Val Asp Glu Cys
            580                 585                 590

Lys Glu Val Pro Asp Ala Cys Phe Asn His Asn Gly Glu His Arg Cys
        595                 600                 605
```

```
Glu Asn Thr Asp Pro Gly Tyr Asn Cys Leu Pro Cys Pro Pro Arg Phe
    610                 615                 620

Thr Gly Ser Gln Pro Phe Gly Gln Gly Val Glu His Ala Thr Ala Asn
625                 630                 635                 640

Lys Gln Val Cys Lys Pro Arg Asn Pro Cys Thr Asp Gly Thr His Asp
                645                 650                 655

Cys Asn Lys Asn Ala Lys Cys Asn Tyr Leu Gly His Tyr Ser Asp Pro
            660                 665                 670

Met Tyr Arg Cys Glu Cys Lys Pro Gly Tyr Ala Gly Asn Gly Ile Ile
        675                 680                 685

Cys Gly Glu Asp Thr Asp Leu Asp Gly Trp Pro Asn Glu Asn Leu Val
    690                 695                 700

Cys Val Ala Asn Ala Thr Tyr His Cys Lys Lys Asp Asn Cys Pro Asn
705                 710                 715                 720

Leu Pro Asn Ser Gly Gln Glu Asp Tyr Asp Lys Asp Gly Ile Gly Asp
                725                 730                 735

Ala Cys Asp Asp Asp Asp Asp Asn Asp Lys Ile Pro Asp Asp Arg Asp
            740                 745                 750

Asn Cys Pro Phe His Tyr Asn Pro Ala Gln Tyr Asp Tyr Asp Arg Asp
        755                 760                 765

Asp Val Gly Asp Arg Cys Asp Asn Cys Pro Tyr Asn His Asn Pro Asp
    770                 775                 780

Gln Ala Asp Thr Asp Asn Asn Gly Glu Gly Asp Ala Cys Ala Ala Asp
785                 790                 795                 800

Ile Asp Gly Asp Gly Ile Leu Asn Glu Arg Asp Asn Cys Gln Tyr Val
                805                 810                 815

Tyr Asn Val Asp Gln Arg Asp Thr Asp Met Asp Gly Val Gly Asp Gln
            820                 825                 830

Cys Asp Asn Cys Pro Leu Glu His Asn Pro Asp Gln Leu Asp Ser Asp
        835                 840                 845
```

```
Ser Asp Arg Ile Gly Asp Thr Cys Asp Asn Asn Gln Asp Ile Asp Glu
    850                 855                 860

Asp Gly His Gln Asn Asn Leu Asp Asn Cys Pro Tyr Val Pro Asn Ala
865                 870                 875                 880

Asn Gln Ala Asp His Asp Lys Asp Gly Lys Gly Asp Ala Cys Asp His
                885                 890                 895

Asp Asp Asp Asn Asp Gly Ile Pro Asp Asp Lys Asp Asn Cys Arg Leu
            900                 905                 910

Val Pro Asn Pro Asp Gln Lys Asp Ser Asp Gly Asp Gly Arg Gly Asp
        915                 920                 925

Ala Cys Lys Asp Asp Phe Asp His Asp Ser Val Pro Asp Ile Asp Asp
    930                 935                 940

Ile Cys Pro Glu Asn Val Asp Ile Ser Glu Thr Asp Phe Arg Arg Phe
945                 950                 955                 960

Gln Met Ile Pro Leu Asp Pro Lys Gly Thr Ser Gln Asn Asp Pro Asn
                965                 970                 975

Trp Val Val Arg His Gln Gly Lys Glu Leu Val Gln Thr Val Asn Cys
            980                 985                 990

Asp Pro Gly Leu Ala Val Gly Tyr Asp Glu Phe Asn Ala Val Asp Phe
        995                 1000                1005

Ser Gly Thr Phe Phe Ile Asn Thr Glu Arg Asp Asp Asp Tyr Ala
    1010                1015                1020

Gly Phe Val Phe Gly Tyr Gln Ser Ser Ser Arg Phe Tyr Val Val
    1025                1030                1035

Met Trp Lys Gln Val Thr Gln Ser Tyr Trp Asp Thr Asn Pro Thr
    1040                1045                1050

Arg Ala Gln Gly Tyr Ser Gly Leu Ser Val Lys Val Val Asn Ser
    1055                1060                1065
```

```
Thr Thr  Gly Pro Gly Glu His  Leu Arg Asn Ala Leu  Trp His Thr
    1070                 1075                 1080

Gly Asn  Thr Pro Gly Gln Val  Arg Thr Leu Trp His  Asp Pro Arg
    1085                 1090                 1095

His Ile  Gly Trp Lys Asp Phe  Thr Ala Tyr Arg Trp  Arg Leu Ser
    1100                 1105                 1110

His Arg  Pro Lys Thr Gly Phe  Ile Arg Val Val Met  Tyr Glu Gly
    1115                 1120                 1125

Lys Lys  Ile Met Ala Asp Ser  Gly Pro Ile Tyr Asp  Lys Thr Tyr
    1130                 1135                 1140

Ala Gly  Gly Arg Leu Gly Leu  Phe Val Phe Ser Gln  Glu Met Val
    1145                 1150                 1155

Phe Phe  Ser Asp Leu Lys Tyr  Glu Cys Arg Asp Pro
    1160                 1165                 1170
```

<210> 31

<211> 380

<212> PRT

<213> Homo sapiens

<400> 31

```
Met Met Phe Ser Gly Phe Asn Ala Asp Tyr Glu Ala Ser Ser Ser Arg
1               5                  10                  15

Cys Ser Ser Ala Ser Pro Ala Gly Asp Ser Leu Ser Tyr Tyr His Ser
            20                  25                  30

Pro Ala Asp Ser Phe Ser Ser Met Gly Ser Pro Val Asn Ala Gln Asp
        35                  40                  45

Phe Cys Thr Asp Leu Ala Val Ser Ser Ala Asn Phe Ile Pro Thr Val
    50                  55                  60

Thr Ala Ile Ser Thr Ser Pro Asp Leu Gln Trp Leu Val Gln Pro Ala
65                  70                  75                  80
```

```
Leu Val Ser Ser Val Ala Pro Ser Gln Thr Arg Ala Pro His Pro Phe
                85                  90                  95

Gly Val Pro Ala Pro Ser Ala Gly Ala Tyr Ser Arg Ala Gly Val Val
            100                 105                 110

Lys Thr Met Thr Gly Gly Arg Ala Gln Ser Ile Gly Arg Arg Gly Lys
        115                 120                 125

Val Glu Gln Leu Ser Pro Glu Glu Glu Glu Lys Arg Arg Ile Arg Arg
    130                 135                 140

Glu Arg Asn Lys Met Ala Ala Ala Lys Cys Arg Asn Arg Arg Arg Glu
145                 150                 155                 160

Leu Thr Asp Thr Leu Gln Ala Glu Thr Asp Gln Leu Glu Asp Glu Lys
                165                 170                 175

Ser Ala Leu Gln Thr Glu Ile Ala Asn Leu Leu Lys Glu Lys Glu Lys
            180                 185                 190

Leu Glu Phe Ile Leu Ala Ala His Arg Pro Ala Cys Lys Ile Pro Asp
        195                 200                 205

Asp Leu Gly Phe Pro Glu Glu Met Ser Val Ala Ser Leu Asp Leu Thr
    210                 215                 220

Gly Gly Leu Pro Glu Val Ala Thr Pro Glu Ser Glu Glu Ala Phe Thr
225                 230                 235                 240

Leu Pro Leu Leu Asn Asp Pro Glu Pro Lys Pro Ser Val Glu Pro Val
                245                 250                 255

Lys Ser Ile Ser Ser Met Glu Leu Lys Thr Glu Pro Phe Asp Asp Phe
            260                 265                 270

Leu Phe Pro Ala Ser Ser Arg Pro Ser Gly Ser Glu Thr Ala Arg Ser
        275                 280                 285

Val Pro Asp Met Asp Leu Ser Gly Ser Phe Tyr Ala Ala Asp Trp Glu
    290                 295                 300

Pro Leu His Ser Gly Ser Leu Gly Met Gly Pro Met Ala Thr Glu Leu
305                 310                 315                 320
```

Glu Pro Leu Cys Thr Pro Val Val Thr Cys Thr Pro Ser Cys Thr Ala
325 330 335

Tyr Thr Ser Ser Phe Val Phe Thr Tyr Pro Glu Ala Asp Ser Phe Pro
340 345 350

Ser Cys Ala Ala Ala His Arg Lys Gly Ser Ser Ser Asn Glu Pro Ser
355 360 365

Ser Asp Ser Leu Ser Ser Pro Thr Leu Leu Ala Leu
370 375 380

<210> 32

<211> 347

<212> PRT

<213> Homo sapiens

<400> 32

Met Cys Thr Lys Met Glu Gln Pro Phe Tyr His Asp Asp Ser Tyr Thr
1 5 10 15

Ala Thr Gly Tyr Gly Arg Ala Pro Gly Gly Leu Ser Leu His Asp Tyr
20 25 30

Lys Leu Leu Lys Pro Ser Leu Ala Val Asn Leu Ala Asp Pro Tyr Arg
35 40 45

Ser Leu Lys Ala Pro Gly Ala Arg Gly Pro Gly Pro Glu Gly Gly Gly
50 55 60

Gly Gly Ser Tyr Phe Ser Gly Gln Gly Ser Asp Thr Gly Ala Ser Leu
65 70 75 80

Lys Leu Ala Ser Ser Glu Leu Glu Arg Leu Ile Val Pro Asn Ser Asn
85 90 95

Gly Val Ile Thr Thr Thr Pro Thr Pro Pro Gly Gln Tyr Phe Tyr Pro
100 105 110

Arg Gly Gly Gly Ser Gly Gly Gly Ala Gly Gly Ala Gly Gly Gly Val
115 120 125

Thr Glu Glu Gln Glu Gly Phe Ala Asp Gly Phe Val Lys Ala Leu Asp
130 135 140

Asp Leu His Lys Met Asn His Val Thr Pro Pro Asn Val Ser Leu Gly
145 150 155 160

Ala Thr Gly Gly Pro Pro Ala Gly Pro Gly Gly Val Tyr Ala Gly Pro
165 170 175

Glu Pro Pro Pro Val Tyr Thr Asn Leu Ser Ser Tyr Ser Pro Ala Ser
180 185 190

Ala Ser Ser Gly Gly Ala Gly Ala Ala Val Gly Thr Gly Ser Ser Tyr
195 200 205

Pro Thr Thr Thr Ile Ser Tyr Leu Pro His Ala Pro Pro Phe Ala Gly
210 215 220

Gly His Pro Ala Gln Leu Gly Leu Gly Arg Gly Ala Ser Thr Phe Lys
225 230 235 240

Glu Glu Pro Gln Thr Val Pro Glu Ala Arg Ser Arg Asp Ala Thr Pro
245 250 255

Pro Val Ser Pro Ile Asn Met Glu Asp Gln Glu Arg Ile Lys Val Glu
260 265 270

Arg Lys Arg Leu Arg Asn Arg Leu Ala Ala Thr Lys Cys Arg Lys Arg
275 280 285

Lys Leu Glu Arg Ile Ala Arg Leu Glu Asp Lys Val Lys Thr Leu Lys
290 295 300

Ala Glu Asn Ala Gly Leu Ser Ser Thr Ala Gly Leu Leu Arg Glu Gln
305 310 315 320

Val Ala Gln Leu Lys Gln Lys Val Met Thr His Val Ser Asn Gly Cys
325 330 335

Gln Leu Leu Leu Gly Val Lys Gly His Ala Phe
340 345

<210> 33

<211> 196

<212> PRT

<213> Homo sapiens

<400> 33

Met Pro Gly Met Phe Phe Ser Ala Asn Pro Lys Glu Leu Lys Gly Thr
1 5 10 15

Thr His Ser Leu Leu Asp Asp Lys Met Gln Lys Arg Arg Pro Lys Thr
20 25 30

Phe Gly Met Asp Met Lys Ala Tyr Leu Arg Ser Met Ile Pro His Leu
35 40 45

Glu Ser Gly Met Lys Ser Ser Lys Ser Lys Asp Val Leu Ser Ala Ala
50 55 60

Glu Val Met Gln Trp Ser Gln Ser Leu Glu Lys Leu Leu Ala Asn Gln
65 70 75 80

Thr Gly Gln Asn Val Phe Gly Ser Phe Leu Lys Ser Glu Phe Ser Glu
85 90 95

Glu Asn Ile Glu Phe Trp Leu Ala Cys Glu Asp Tyr Lys Lys Thr Glu
100 105 110

Ser Asp Leu Leu Pro Cys Lys Ala Glu Glu Ile Tyr Lys Ala Phe Val
115 120 125

His Ser Asp Ala Ala Lys Gln Ile Asn Ile Asp Phe Arg Thr Arg Glu
130 135 140

Ser Thr Ala Lys Lys Ile Lys Ala Pro Thr Pro Thr Cys Phe Asp Glu
145 150 155 160

Ala Gln Lys Val Ile Tyr Thr Leu Met Glu Lys Asp Ser Tyr Pro Arg
165 170 175

Phe Leu Lys Ser Asp Ile Tyr Leu Asn Leu Leu Asn Asp Leu Gln Ala
180 185 190

Asn Ser Leu Lys
195

<210> 34

<211> 367

<212> PRT

<213> Homo sapiens

<400> 34

Met Val Met Glu Val Gly Thr Leu Asp Ala Gly Gly Leu Arg Ala Leu
1 5 10 15

Leu Gly Glu Arg Ala Ala Gln Cys Leu Leu Leu Asp Cys Arg Ser Phe
20 25 30

Phe Ala Phe Asn Ala Gly His Ile Ala Gly Ser Val Asn Val Arg Phe
35 40 45

Ser Thr Ile Val Arg Arg Arg Ala Lys Gly Ala Met Gly Leu Glu His
50 55 60

Ile Val Pro Asn Ala Glu Leu Arg Gly Arg Leu Leu Ala Gly Ala Tyr
65 70 75 80

His Ala Val Val Leu Leu Asp Glu Arg Ser Ala Ala Leu Asp Gly Ala
85 90 95

Lys Arg Asp Gly Thr Leu Ala Leu Ala Ala Gly Ala Leu Cys Arg Glu
100 105 110

Ala Arg Ala Ala Gln Val Phe Phe Leu Lys Gly Gly Tyr Glu Ala Phe
115 120 125

Ser Ala Ser Cys Pro Glu Leu Cys Ser Lys Gln Ser Thr Pro Met Gly
130 135 140

Leu Ser Leu Pro Leu Ser Thr Ser Val Pro Asp Ser Ala Glu Ser Gly
145 150 155 160

```
Cys Ser Ser Cys Ser Thr Pro Leu Tyr Asp Gln Gly Gly Pro Val Glu
                165                 170                 175

Ile Leu Pro Phe Leu Tyr Leu Gly Ser Ala Tyr His Ala Ser Arg Lys
            180                 185                 190

Asp Met Leu Asp Ala Leu Gly Ile Thr Ala Leu Ile Asn Val Ser Ala
        195                 200                 205

Asn Cys Pro Asn His Phe Glu Gly His Tyr Gln Tyr Lys Ser Ile Pro
    210                 215                 220

Val Glu Asp Asn His Lys Ala Asp Ile Ser Ser Trp Phe Asn Glu Ala
225                 230                 235                 240

Ile Asp Phe Ile Asp Ser Ile Lys Asn Ala Gly Gly Arg Val Phe Val
                245                 250                 255

His Cys Gln Ala Gly Ile Ser Arg Ser Ala Thr Ile Cys Leu Ala Tyr
            260                 265                 270

Leu Met Arg Thr Asn Arg Val Lys Leu Asp Glu Ala Phe Glu Phe Val
        275                 280                 285

Lys Gln Arg Arg Ser Ile Ile Ser Pro Asn Phe Ser Phe Met Gly Gln
    290                 295                 300

Leu Leu Gln Phe Glu Ser Gln Val Leu Ala Pro His Cys Ser Ala Glu
305                 310                 315                 320

Ala Gly Ser Pro Ala Met Ala Val Leu Asp Arg Gly Thr Ser Thr Thr
                325                 330                 335

Thr Val Phe Asn Phe Pro Val Ser Ile Pro Val His Ser Thr Asn Ser
            340                 345                 350

Ala Leu Ser Tyr Leu Gln Ser Pro Ile Thr Thr Ser Pro Ser Cys
        355                 360                 365
```

<210> 35

<211> 107

<212> PRT

<213> Homo sapiens

<400> 35

Met Ala Arg Ala Thr Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu
1 5 10 15

Arg Val Ala Leu Leu Leu Leu Leu Leu Val Ala Ala Ser Arg Arg Ala
20 25 30

Ala Gly Ala Pro Leu Ala Thr Glu Leu Arg Cys Gln Cys Leu Gln Thr
35 40 45

Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val Lys Val Lys Ser
50 55 60

Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn
65 70 75 80

Gly Gln Lys Ala Cys Leu Asn Pro Ala Ser Pro Met Val Lys Lys Ile
85 90 95

Ile Glu Lys Met Leu Lys Asn Gly Lys Ser Asn
100 105

<210> 36

<211> 106

<212> PRT

<213> Homo sapiens

<400> 36

Met Ala His Ala Thr Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu
1 5 10 15

Arg Val Ala Leu Leu Leu Leu Leu Leu Val Gly Ser Arg Arg Ala Ala
20 25 30

Gly Ala Ser Val Val Thr Glu Leu Arg Cys Gln Cys Leu Gln Thr Leu
35 40 45

Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val Asn Val Arg Ser Pro
50 55 60

Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn Gly
65 70 75 80

Lys Lys Ala Cys Leu Asn Pro Ala Ser Pro Met Val Gln Lys Ile Ile
85 90 95

Glu Lys Ile Leu Asn Lys Gly Ser Thr Asn
100 105

<210> 37

<211> 98

<212> PRT

<213> Homo sapiens

<400> 37

Met Asn Gln Thr Ala Ile Leu Ile Cys Cys Leu Ile Phe Leu Thr Leu
1 5 10 15

Ser Gly Ile Gln Gly Val Pro Leu Ser Arg Thr Val Arg Cys Thr Cys
20 25 30

Ile Ser Ile Ser Asn Gln Pro Val Asn Pro Arg Ser Leu Glu Lys Leu
35 40 45

Glu Ile Ile Pro Ala Ser Gln Phe Cys Pro Arg Val Glu Ile Ile Ala
50 55 60

Thr Met Lys Lys Lys Gly Glu Lys Arg Cys Leu Asn Pro Glu Ser Lys
65 70 75 80

Ala Ile Lys Asn Leu Leu Lys Ala Val Ser Lys Glu Met Ser Lys Arg
85 90 95

Ser Pro

<210> 38

<211> 564

<212> PRT

<213> Homo sapiens

<400> 38

```
Met Lys Glu His Gly Gly Thr Phe Ser Ser Thr Gly Ile Ser Gly Gly
1               5                  10                  15

Ser Gly Asp Ser Ala Met Asp Ser Leu Gln Pro Leu Gln Pro Asn Tyr
            20                  25                  30

Met Pro Val Cys Leu Phe Ala Glu Glu Ser Tyr Gln Lys Leu Ala Met
        35                  40                  45

Glu Thr Leu Glu Glu Leu Asp Trp Cys Leu Asp Gln Leu Glu Thr Ile
    50                  55                  60

Gln Thr Tyr Arg Ser Val Ser Glu Met Ala Ser Asn Lys Phe Lys Arg
65                  70                  75                  80

Met Leu Asn Arg Glu Leu Thr His Leu Ser Glu Met Ser Arg Ser Gly
                85                  90                  95

Asn Gln Val Ser Glu Tyr Ile Ser Asn Thr Phe Leu Asp Lys Gln Asn
            100                 105                 110

Asp Val Glu Ile Pro Ser Pro Thr Gln Lys Asp Arg Glu Lys Lys Lys
        115                 120                 125

Lys Gln Gln Leu Met Thr Gln Ile Ser Gly Val Lys Lys Leu Met His
    130                 135                 140

Ser Ser Ser Leu Asn Asn Thr Ser Ile Ser Arg Phe Gly Val Asn Thr
145                 150                 155                 160

Glu Asn Glu Asp His Leu Ala Lys Glu Leu Glu Asp Leu Asn Lys Trp
                165                 170                 175

Gly Leu Asn Ile Phe Asn Val Ala Gly Tyr Ser His Asn Arg Pro Leu
            180                 185                 190
```

Thr Cys Ile Met Tyr Ala Ile Phe Gln Glu Arg Asp Leu Leu Lys Thr
195 200 205

Phe Arg Ile Ser Ser Asp Thr Phe Ile Thr Tyr Met Met Thr Leu Glu
210 215 220

Asp His Tyr His Ser Asp Val Ala Tyr His Asn Ser Leu His Ala Ala
225 230 235 240

Asp Val Ala Gln Ser Thr His Val Leu Leu Ser Thr Pro Ala Leu Asp
245 250 255

Ala Val Phe Thr Asp Leu Glu Ile Leu Ala Ala Ile Phe Ala Ala Ala
260 265 270

Ile His Asp Val Asp His Pro Gly Val Ser Asn Gln Phe Leu Ile Asn
275 280 285

Thr Asn Ser Glu Leu Ala Leu Met Tyr Asn Asp Glu Ser Val Leu Glu
290 295 300

Asn His His Leu Ala Val Gly Phe Lys Leu Leu Gln Glu Glu His Cys
305 310 315 320

Asp Ile Phe Met Asn Leu Thr Lys Lys Gln Arg Gln Thr Leu Arg Lys
325 330 335

Met Val Ile Asp Met Val Leu Ala Thr Asp Met Ser Lys His Met Ser
340 345 350

Leu Leu Ala Asp Leu Lys Thr Met Val Glu Thr Lys Lys Val Thr Ser
355 360 365

Ser Gly Val Leu Leu Leu Asp Asn Tyr Thr Asp Arg Ile Gln Val Leu
370 375 380

Arg Asn Met Val His Cys Ala Asp Leu Ser Asn Pro Thr Lys Ser Leu
385 390 395 400

Glu Leu Tyr Arg Gln Trp Thr Asp Arg Ile Met Glu Glu Phe Phe Gln
405 410 415

Gln Gly Asp Lys Glu Arg Glu Arg Gly Met Glu Ile Ser Pro Met Cys
420 425 430

```
Asp Lys His Thr Ala Ser Val Glu Lys Ser Gln Val Gly Phe Ile Asp
         435                 440                 445

Tyr Ile Val His Pro Leu Trp Glu Thr Trp Ala Asp Leu Val Gln Pro
     450                 455                 460

Asp Ala Gln Asp Ile Leu Asp Thr Leu Glu Asp Asn Arg Asn Trp Tyr
465                 470                 475                 480

Gln Ser Met Ile Pro Gln Ser Pro Ser Pro Pro Leu Asp Glu Gln Asn
                485                 490                 495

Arg Asp Cys Gln Gly Leu Met Glu Lys Phe Gln Phe Glu Leu Thr Leu
            500                 505                 510

Asp Glu Glu Asp Ser Glu Gly Pro Glu Lys Glu Gly Glu Gly His Ser
         515                 520                 525

Tyr Phe Ser Ser Thr Lys Thr Leu Cys Val Ile Asp Pro Glu Asn Arg
     530                 535                 540

Asp Ser Leu Gly Glu Thr Asp Ile Asp Ile Ala Thr Glu Asp Lys Ser
545                 550                 555                 560

Pro Val Asp Thr
```

<210> 39

<211> 326

<212> PRT

<213> Homo sapiens

<400> 39

```
Met Pro Pro Ser Ile Ser Ala Phe Gln Ala Ala Tyr Ile Gly Ile Glu
1               5                   10                  15

Val Leu Ile Ala Leu Val Ser Val Pro Gly Asn Val Leu Val Ile Trp
            20                  25                  30
```

```
Ala Val Lys Val Asn Gln Ala Leu Arg Asp Ala Thr Phe Cys Phe Ile
        35                  40                  45

Val Ser Leu Ala Val Ala Asp Val Ala Val Gly Ala Leu Val Ile Pro
    50                  55                  60

Leu Ala Ile Leu Ile Asn Ile Gly Pro Gln Thr Tyr Phe His Thr Cys
65                  70                  75                  80

Leu Met Val Ala Cys Pro Val Leu Ile Leu Thr Gln Ser Ser Ile Leu
                85                  90                  95

Ala Leu Leu Ala Ile Ala Val Asp Arg Tyr Leu Arg Val Lys Ile Pro
            100                 105                 110

Leu Arg Tyr Lys Met Val Val Thr Pro Arg Arg Ala Ala Val Ala Ile
        115                 120                 125

Ala Gly Cys Trp Ile Leu Ser Phe Val Val Gly Leu Thr Pro Met Phe
    130                 135                 140

Gly Trp Asn Asn Leu Ser Ala Val Glu Arg Ala Trp Ala Ala Asn Gly
145                 150                 155                 160

Ser Met Gly Glu Pro Val Ile Lys Cys Glu Phe Glu Lys Val Ile Ser
                165                 170                 175

Met Glu Tyr Met Val Tyr Phe Asn Phe Phe Val Trp Val Leu Pro Pro
            180                 185                 190

Leu Leu Leu Met Val Leu Ile Tyr Leu Glu Val Phe Tyr Leu Ile Arg
        195                 200                 205

Lys Gln Leu Asn Lys Lys Val Ser Ala Ser Ser Gly Asp Pro Gln Lys
    210                 215                 220

Tyr Tyr Gly Lys Glu Leu Lys Ile Ala Lys Ser Leu Ala Leu Ile Leu
225                 230                 235                 240

Phe Leu Phe Ala Leu Ser Trp Leu Pro Leu His Ile Leu Asn Cys Ile
                245                 250                 255

Thr Leu Phe Cys Pro Ser Cys His Lys Pro Ser Ile Leu Thr Tyr Ile
            260                 265                 270
```

Ala Ile Phe Leu Thr His Gly Asn Ser Ala Met Asn Pro Ile Val Tyr
275 280 285

Ala Phe Arg Ile Gln Lys Phe Arg Val Thr Phe Leu Lys Ile Trp Asn
290 295 300

Asp His Phe Arg Cys Gln Pro Ala Pro Pro Ile Asp Glu Asp Leu Pro
305 310 315 320

Glu Glu Arg Pro Asp Asp
325

<210> 40

<211> 302

<212> PRT

<213> Homo sapiens

<400> 40

Ala Asn Ser Val Val Val Trp Val Asn Ile Gln Ala Lys Thr Thr Gly
1 5 10 15

Tyr Asp Thr His Cys Tyr Ile Leu Asn Leu Ala Ile Ala Asp Leu Trp
20 25 30

Val Val Leu Thr Ile Pro Val Trp Val Val Ser Leu Val Gln His Asn
35 40 45

Gln Trp Pro Met Gly Glu Leu Thr Cys Lys Val Thr His Leu Ile Phe
50 55 60

Ser Ile Asn Leu Phe Ser Ser Ile Phe Phe Leu Thr Cys Met Ser Val
65 70 75 80

Asp Arg Tyr Leu Ser Ile Thr Tyr Phe Thr Asn Thr Pro Ser Ser Arg
85 90 95

Lys Lys Met Val Arg Arg Val Val Cys Ile Leu Val Trp Leu Leu Ala
100 105 110

Phe Cys Val Ser Leu Pro Asp Thr Tyr Tyr Leu Lys Thr Val Thr Ser
115 120 125

```
Ala Ser Asn Asn Glu Thr Tyr Cys Arg Ser Phe Tyr Pro Glu His Ser
    130                 135                 140

Ile Lys Glu Trp Leu Ile Gly Met Glu Leu Val Ser Val Val Leu Gly
145                 150                 155                 160

Phe Ala Val Pro Phe Ser Ile Ile Ala Val Phe Tyr Phe Leu Leu Ala
                165                 170                 175

Arg Ala Ile Ser Ala Ser Ser Asp Gln Glu Lys His Ser Ser Arg Lys
            180                 185                 190

Ile Ile Phe Ser Tyr Val Val Val Phe Leu Val Cys Trp Leu Pro Tyr
        195                 200                 205

His Val Ala Val Leu Leu Asp Ile Phe Ser Ile Leu His Tyr Ile Pro
    210                 215                 220

Phe Thr Cys Arg Leu Glu His Ala Leu Phe Thr Ala Leu His Val Thr
225                 230                 235                 240

Gln Cys Leu Ser Leu Val His Cys Cys Val Asn Pro Val Leu Tyr Ser
                245                 250                 255

Phe Ile Asn Arg Asn Tyr Arg Tyr Glu Leu Met Lys Ala Phe Ile Phe
            260                 265                 270

Lys Tyr Ser Ala Lys Thr Gly Leu Thr Lys Leu Ile Asp Ala Ser Arg
        275                 280                 285

Val Ser Glu Thr Glu Tyr Ser Ala Leu Glu Gln Ser Thr Lys
    290                 295                 300
```

<210> 41

<211> 380

<212> PRT

<213> Homo sapiens

<400> 41

```
Met Met Phe Ser Gly Phe Asn Ala Asp Tyr Glu Ala Ser Ser Ser Arg
1               5                  10                  15

Cys Ser Ser Ala Ser Pro Ala Gly Asp Ser Leu Ser Tyr Tyr His Ser
            20                  25                  30

Pro Ala Asp Ser Phe Ser Ser Met Gly Ser Pro Val Asn Ala Gln Asp
        35                  40                  45

Phe Cys Thr Asp Leu Ala Val Ser Ser Ala Asn Phe Ile Pro Thr Val
    50                  55                  60

Thr Ala Ile Ser Thr Ser Pro Asp Leu Gln Trp Leu Val Gln Pro Ala
65                  70                  75                  80

Leu Val Ser Ser Val Ala Pro Ser Gln Thr Arg Ala Pro His Pro Phe
                85                  90                  95

Gly Val Pro Ala Pro Ser Ala Gly Ala Tyr Ser Arg Ala Gly Val Val
            100                 105                 110

Lys Thr Met Thr Gly Gly Arg Ala Gln Ser Ile Gly Arg Arg Gly Lys
        115                 120                 125

Val Glu Gln Leu Ser Pro Glu Glu Glu Glu Lys Arg Arg Ile Arg Arg
    130                 135                 140

Glu Arg Asn Lys Met Ala Ala Ala Lys Cys Arg Asn Arg Arg Arg Glu
145                 150                 155                 160

Leu Thr Asp Thr Leu Gln Ala Glu Thr Asp Gln Leu Glu Asp Glu Lys
                165                 170                 175

Ser Ala Leu Gln Thr Glu Ile Ala Asn Leu Leu Lys Glu Lys Glu Lys
            180                 185                 190

Leu Glu Phe Ile Leu Ala Ala His Arg Pro Ala Cys Lys Ile Pro Asp
        195                 200                 205

Asp Leu Gly Phe Pro Glu Glu Met Ser Val Ala Ser Leu Asp Leu Thr
    210                 215                 220
```

Gly Gly Leu Pro Glu Val Ala Thr Pro Glu Ser Glu Glu Ala Phe Thr
225 230 235 240

Leu Pro Leu Leu Asn Asp Pro Glu Pro Lys Pro Ser Val Glu Pro Val
245 250 255

Lys Ser Ile Ser Ser Met Glu Leu Lys Thr Glu Pro Phe Asp Asp Phe
260 265 270

Leu Phe Pro Ala Ser Ser Arg Pro Ser Gly Ser Glu Thr Ala Arg Ser
275 280 285

Val Pro Asp Met Asp Leu Ser Gly Ser Phe Tyr Ala Ala Asp Trp Glu
290 295 300

Pro Leu His Ser Gly Ser Leu Gly Met Gly Pro Met Ala Thr Glu Leu
305 310 315 320

Glu Pro Leu Cys Thr Pro Val Val Thr Cys Thr Pro Ser Cys Thr Ala
325 330 335

Tyr Thr Ser Ser Phe Val Phe Thr Tyr Pro Glu Ala Asp Ser Phe Pro
340 345 350

Ser Cys Ala Ala Ala His Arg Lys Gly Ser Ser Ser Asn Glu Pro Ser
355 360 365

Ser Asp Ser Leu Ser Ser Pro Thr Leu Leu Ala Leu
370 375 380

<210> 42

<211> 119

<212> PRT

<213> Homo sapiens

<400> 42

Met Lys Ala Leu Ser Pro Val Arg Gly Cys Tyr Glu Ala Val Cys Cys
1 5 10 15

Leu Ser Glu Arg Ser Leu Ala Ile Ala Arg Gly Arg Gly Lys Gly Pro
20 25 30

Ala Ala Glu Glu Pro Leu Ser Leu Leu Asp Asp Met Asn His Cys Tyr
35 40 45

Ser Arg Leu Arg Glu Leu Val Pro Gly Val Pro Arg Gly Thr Gln Leu
50 55 60

Ser Gln Val Glu Ile Leu Gln Arg Val Ile Asp Tyr Ile Leu Asp Leu
65 70 75 80

Gln Val Val Leu Ala Glu Pro Ala Pro Gly Pro Pro Asp Gly Pro His
85 90 95

Leu Pro Ile Gln Thr Ala Glu Leu Thr Pro Glu Leu Val Ile Ser Asn
100 105 110

Asp Lys Arg Ser Phe Cys His
115

<210> 43

<211> 496

<212> PRT

<213> Homo sapiens

<400> 43

Met Gly Thr Gln Lys Val Thr Pro Ala Leu Ile Phe Ala Ile Thr Val
1 5 10 15

Ala Thr Ile Gly Ser Phe Gln Phe Gly Tyr Asn Thr Gly Val Ile Asn
20 25 30

Ala Pro Glu Lys Ile Ile Lys Glu Phe Ile Asn Lys Thr Leu Thr Asp
35 40 45

Lys Gly Asn Ala Pro Pro Ser Glu Val Leu Leu Thr Ser Leu Trp Ser
50 55 60

Leu Ser Val Ala Ile Phe Ser Val Gly Gly Met Ile Gly Ser Phe Ser
65 70 75 80

```
Val Gly Leu Phe Val Asn Arg Phe Gly Arg Arg Asn Ser Met Leu Ile
                85                  90                  95

Val Asn Leu Leu Ala Val Thr Gly Gly Cys Phe Met Gly Leu Cys Lys
            100                 105                 110

Val Ala Lys Ser Val Glu Met Leu Ile Leu Gly Arg Leu Val Ile Gly
        115                 120                 125

Leu Phe Cys Gly Leu Cys Thr Gly Phe Val Pro Met Tyr Ile Gly Glu
    130                 135                 140

Ile Ser Pro Thr Ala Leu Arg Gly Ala Phe Gly Thr Leu Asn Gln Leu
145                 150                 155                 160

Gly Ile Val Val Gly Ile Leu Val Ala Gln Ile Phe Gly Leu Glu Phe
                165                 170                 175

Ile Leu Gly Ser Glu Glu Leu Trp Pro Leu Leu Leu Gly Phe Thr Ile
            180                 185                 190

Leu Pro Ala Ile Leu Gln Ser Ala Ala Leu Pro Phe Cys Pro Glu Ser
        195                 200                 205

Pro Arg Phe Leu Leu Ile Asn Arg Lys Glu Glu Glu Asn Ala Lys Gln
    210                 215                 220

Ile Leu Gln Arg Leu Trp Gly Thr Gln Asp Val Ser Gln Asp Ile Gln
225                 230                 235                 240

Glu Met Lys Asp Glu Ser Ala Arg Met Ser Gln Glu Lys Gln Val Thr
                245                 250                 255

Val Leu Glu Leu Phe Arg Val Ser Ser Tyr Arg Gln Pro Ile Ile Ile
            260                 265                 270

Ser Ile Val Leu Gln Leu Ser Gln Gln Leu Ser Gly Ile Asn Ala Val
        275                 280                 285

Phe Tyr Tyr Ser Thr Gly Ile Phe Lys Asp Ala Gly Val Gln Glu Pro
    290                 295                 300

Ile Tyr Ala Thr Ile Gly Ala Gly Val Val Asn Thr Ile Phe Thr Val
305                 310                 315                 320
```

Val Ser Leu Phe Leu Val Glu Arg Ala Gly Arg Arg Thr Leu His Met
325 330 335

Ile Gly Leu Gly Gly Met Ala Phe Cys Ser Thr Leu Met Thr Val Ser
340 345 350

Leu Leu Leu Lys Asp Asn Tyr Asn Gly Met Ser Phe Val Cys Ile Gly
355 360 365

Ala Ile Leu Val Phe Val Ala Phe Phe Glu Ile Gly Pro Gly Pro Ile
370 375 380

Pro Trp Phe Ile Val Ala Glu Leu Phe Ser Gln Gly Pro Arg Pro Ala
385 390 395 400

Ala Met Ala Val Ala Gly Cys Ser Asn Trp Thr Ser Asn Phe Leu Val
405 410 415

Gly Leu Leu Phe Pro Ser Ala Ala His Tyr Leu Gly Ala Tyr Val Phe
420 425 430

Ile Ile Phe Thr Gly Phe Leu Ile Thr Phe Leu Ala Phe Thr Phe Phe
435 440 445

Lys Val Pro Glu Thr Arg Gly Arg Thr Phe Glu Asp Ile Thr Arg Ala
450 455 460

Phe Glu Gly Gln Ala His Gly Ala Asp Arg Ser Gly Lys Asp Gly Val
465 470 475 480

Met Glu Met Asn Ser Ile Glu Pro Ala Lys Glu Thr Thr Thr Asn Val
485 490 495

<210> 44

<211> 352

<212> PRT

<213> Homo sapiens

<400> 44

```
Met Glu Gly Ile Ser Ile Tyr Thr Ser Asp Asn Tyr Thr Glu Glu Met
1               5                  10                  15

Gly Ser Gly Asp Tyr Asp Ser Met Lys Glu Pro Cys Phe Arg Glu Glu
            20                  25                  30

Asn Ala Asn Phe Asn Lys Ile Phe Leu Pro Thr Ile Tyr Ser Ile Ile
        35                  40                  45

Phe Leu Thr Gly Ile Val Gly Asn Gly Leu Val Ile Leu Val Met Gly
    50                  55                  60

Tyr Gln Lys Lys Leu Arg Ser Met Thr Asp Lys Tyr Arg Leu His Leu
65                  70                  75                  80

Ser Val Ala Asp Leu Leu Phe Val Ile Thr Leu Pro Phe Trp Ala Val
                85                  90                  95

Asp Ala Val Ala Asn Trp Tyr Phe Gly Asn Phe Leu Cys Lys Ala Val
            100                 105                 110

His Val Ile Tyr Thr Val Asn Leu Tyr Ser Ser Val Leu Ile Leu Ala
        115                 120                 125

Phe Ile Ser Leu Asp Arg Tyr Leu Ala Ile Val His Ala Thr Asn Ser
    130                 135                 140

Gln Arg Pro Arg Lys Leu Leu Ala Glu Lys Val Val Tyr Val Gly Val
145                 150                 155                 160

Trp Ile Pro Ala Leu Leu Leu Thr Ile Pro Asp Phe Ile Phe Ala Asn
                165                 170                 175

Val Ser Glu Ala Asp Asp Arg Tyr Ile Cys Asp Arg Phe Tyr Pro Asn
            180                 185                 190

Asp Leu Trp Val Val Val Phe Gln Phe Gln His Ile Met Val Gly Leu
        195                 200                 205

Ile Leu Pro Gly Ile Val Ile Leu Ser Cys Tyr Cys Ile Ile Ile Ser
    210                 215                 220
```

```
Lys Leu Ser His Ser Lys Gly His Gln Lys Arg Lys Ala Leu Lys Thr
225                 230                 235                 240

Thr Val Ile Leu Ile Leu Ala Phe Phe Ala Cys Trp Leu Pro Tyr Tyr
                245                 250                 255

Ile Gly Ile Ser Ile Asp Ser Phe Ile Leu Leu Glu Ile Ile Lys Gln
            260                 265                 270

Gly Cys Glu Phe Glu Asn Thr Val His Lys Trp Ile Ser Ile Thr Glu
        275                 280                 285

Ala Leu Ala Phe Phe His Cys Cys Leu Asn Pro Ile Leu Tyr Ala Phe
    290                 295                 300

Leu Gly Ala Lys Phe Lys Thr Ser Ala Gln His Ala Leu Thr Ser Val
305                 310                 315                 320

Ser Arg Gly Ser Ser Leu Lys Ile Leu Ser Lys Gly Lys Arg Gly Gly
                325                 330                 335

His Ser Ser Val Ser Thr Glu Ser Glu Ser Ser Ser Phe His Ser Ser
            340                 345                 350
```

<210> 45

<211> 208

<212> PRT

<213> Homo sapiens

<400> 45

```
Met Lys Leu Leu Pro Ser Val Val Leu Lys Leu Phe Leu Ala Ala Val
1               5                   10                  15

Leu Ser Ala Leu Val Thr Gly Glu Ser Leu Glu Arg Leu Arg Arg Gly
            20                  25                  30

Leu Ala Ala Gly Thr Ser Asn Pro Asp Pro Pro Thr Val Ser Thr Asp
        35                  40                  45

Gln Leu Leu Pro Leu Gly Gly Gly Arg Asp Arg Lys Val Arg Asp Leu
    50                  55                  60
```

Gln Glu Ala Asp Leu Asp Leu Leu Arg Val Thr Leu Ser Ser Lys Pro
65 70 75 80

Gln Ala Leu Ala Thr Pro Asn Lys Glu Glu His Gly Lys Arg Lys Lys
85 90 95

Lys Gly Lys Gly Leu Gly Lys Lys Arg Asp Pro Cys Leu Arg Lys Tyr
100 105 110

Lys Asp Phe Cys Ile His Gly Glu Cys Lys Tyr Val Lys Glu Leu Arg
115 120 125

Ala Pro Ser Cys Ile Cys His Pro Gly Tyr His Gly Glu Arg Cys His
130 135 140

Gly Leu Ser Leu Pro Val Glu Asn Arg Leu Tyr Thr Tyr Asp His Thr
145 150 155 160

Thr Ile Leu Ala Val Val Ala Val Val Leu Ser Ser Val Cys Leu Leu
165 170 175

Val Ile Val Gly Leu Leu Met Phe Arg Tyr His Arg Arg Gly Gly Tyr
180 185 190

Asp Val Glu Asn Glu Glu Lys Val Lys Leu Gly Met Thr Asn Ser His
195 200 205

<210> 46

<211> 218

<212> PRT

<213> Homo sapiens

<400> 46

Met Ser Met Thr Leu Gly Tyr Trp Asp Ile Arg Gly Leu Ala His Ala
1 5 10 15

Met Arg Leu Leu Leu Glu Tyr Thr Asp Ser Ser Tyr Glu Glu Lys Lys
20 25 30

Tyr Thr Met Gly Asp Ala Pro Asp Tyr Asp Arg Ser Gln Trp Leu Asn
35 40 45

```
Glu Lys Phe Lys Leu Gly Leu Asp Phe Pro Asn Leu Pro Tyr Leu Ile
     50                  55                  60

Asp Gly Ala His Lys Ile Thr Gln Ser Asn Ala Ile Leu Cys Tyr Ile
65                  70                  75                  80

Ala Arg Lys His Asn Leu Cys Gly Glu Thr Glu Glu Glu Lys Ile Arg
                85                  90                  95

Val Asp Ile Leu Glu Asn Gln Ala Met Asp Val Ser Asn Gln Leu Ala
            100                 105                 110

Arg Val Cys Tyr Ser Pro Asp Phe Glu Lys Leu Lys Pro Glu Tyr Leu
        115                 120                 125

Glu Glu Leu Pro Thr Met Met Gln His Phe Ser Gln Phe Leu Gly Lys
    130                 135                 140

Arg Pro Trp Phe Val Gly Asp Lys Ile Thr Phe Val Asp Phe Leu Ala
145                 150                 155                 160

Tyr Asp Val Leu Asp Leu His Arg Ile Phe Glu Pro Asn Cys Leu Asp
                165                 170                 175

Ala Phe Pro Asn Leu Lys Asp Phe Ile Ser Arg Phe Glu Gly Leu Glu
            180                 185                 190

Lys Ile Ser Ala Tyr Met Lys Ser Ser Arg Phe Leu Pro Lys Pro Leu
        195                 200                 205

Tyr Thr Arg Val Ala Val Trp Gly Asn Lys
    210                 215
```

<210> 47

<211> 543

<212> PRT

<213> Homo sapiens

<400> 47

```
Met Gly Thr Ser Leu Ser Pro Asn Asp Pro Trp Pro Leu Asn Pro Leu
1               5                  10                  15

Ser Ile Gln Gln Thr Thr Leu Leu Leu Leu Leu Ser Val Leu Ala Thr
            20                  25                  30

Val His Val Gly Gln Arg Leu Leu Arg Gln Arg Arg Arg Gln Leu Arg
        35                  40                  45

Ser Ala Pro Pro Gly Pro Phe Ala Trp Pro Leu Ile Gly Asn Ala Ala
    50                  55                  60

Ala Val Gly Gln Ala Ala His Leu Ser Phe Ala Arg Leu Ala Arg Arg
65                  70                  75                  80

Tyr Gly Asp Val Phe Gln Ile Arg Leu Gly Ser Cys Pro Ile Val Val
                85                  90                  95

Leu Asn Gly Glu Arg Ala Ile His Gln Ala Leu Val Gln Gln Gly Ser
            100                 105                 110

Ala Phe Ala Asp Arg Pro Ala Phe Ala Ser Phe Arg Val Val Ser Gly
        115                 120                 125

Gly Arg Ser Met Ala Phe Gly His Tyr Ser Glu His Trp Lys Val Gln
    130                 135                 140

Arg Arg Ala Ala His Ser Met Met Arg Asn Phe Phe Thr Arg Gln Pro
145                 150                 155                 160

Arg Ser Arg Gln Val Leu Glu Gly His Val Leu Ser Glu Ala Arg Glu
                165                 170                 175

Leu Val Ala Leu Leu Val Arg Gly Ser Ala Asp Gly Ala Phe Leu Asp
            180                 185                 190

Pro Arg Pro Leu Thr Val Val Ala Val Ala Asn Val Met Ser Ala Val
        195                 200                 205

Cys Phe Gly Cys Arg Tyr Ser His Asp Asp Pro Glu Phe Arg Glu Leu
    210                 215                 220
```

```
Leu Ser His Asn Glu Glu Phe Gly Arg Thr Val Gly Ala Gly Ser Leu
225                 230                 235                 240

Val Asp Val Met Pro Trp Leu Gln Tyr Phe Pro Asn Pro Val Arg Thr
                245                 250                 255

Val Phe Arg Glu Phe Glu Gln Leu Asn Arg Asn Phe Ser Asn Phe Ile
            260                 265                 270

Leu Asp Lys Phe Leu Arg His Cys Glu Ser Leu Arg Pro Gly Ala Ala
        275                 280                 285

Pro Arg Asp Met Met Asp Ala Phe Ile Leu Ser Ala Glu Lys Lys Ala
    290                 295                 300

Ala Gly Asp Ser His Gly Gly Gly Ala Arg Leu Asp Leu Glu Asn Val
305                 310                 315                 320

Pro Ala Thr Ile Thr Asp Ile Phe Gly Ala Ser Gln Asp Thr Leu Ser
                325                 330                 335

Thr Ala Leu Gln Trp Leu Leu Leu Leu Phe Thr Arg Tyr Pro Asp Val
            340                 345                 350

Gln Thr Arg Val Gln Ala Glu Leu Asp Gln Val Val Gly Arg Asp Arg
        355                 360                 365

Leu Pro Cys Met Gly Asp Gln Pro Asn Leu Pro Tyr Val Leu Ala Phe
    370                 375                 380

Leu Tyr Glu Ala Met Arg Phe Ser Ser Phe Val Pro Val Thr Ile Pro
385                 390                 395                 400

His Ala Thr Thr Ala Asn Thr Ser Val Leu Gly Tyr His Ile Pro Lys
                405                 410                 415

Asp Thr Val Val Phe Val Asn Gln Trp Ser Val Asn His Asp Pro Val
            420                 425                 430

Lys Trp Pro Asn Pro Glu Asn Phe Asp Pro Ala Arg Phe Leu Asp Lys
        435                 440                 445

Asp Gly Leu Ile Asn Lys Asp Leu Thr Ser Arg Val Met Ile Phe Ser
    450                 455                 460
```

Val Gly Lys Arg Arg Cys Ile Gly Glu Glu Leu Ser Lys Met Gln Leu
465 470 475 480

Phe Leu Phe Ile Ser Ile Leu Ala His Gln Cys Asp Phe Arg Ala Asn
485 490 495

Pro Asn Glu Pro Ala Lys Met Asn Phe Ser Tyr Gly Leu Thr Ile Lys
500 505 510

Pro Lys Ser Phe Lys Val Asn Val Thr Leu Arg Glu Ser Met Glu Leu
515 520 525

Leu Asp Ser Ala Val Gln Asn Leu Gln Ala Lys Glu Thr Cys Gln
530 535 540

<210> 48

<211> 300

<212> PRT

<213> Homo sapiens

<400> 48

Met Arg Ile Ala Val Ile Cys Phe Cys Leu Leu Gly Ile Thr Cys Ala
1 5 10 15

Ile Pro Val Lys Gln Ala Asp Ser Gly Ser Ser Glu Glu Lys Gln Leu
20 25 30

Tyr Asn Lys Tyr Pro Asp Ala Val Ala Thr Trp Leu Asn Pro Asp Pro
35 40 45

Ser Gln Lys Gln Asn Leu Leu Ala Pro Gln Thr Leu Pro Ser Lys Ser
50 55 60

Asn Glu Ser His Asp His Met Asp Asp Met Asp Asp Glu Asp Asp Asp
65 70 75 80

Asp His Val Asp Ser Gln Asp Ser Ile Asp Ser Asn Asp Ser Asp Asp
85 90 95

Val Asp Asp Thr Asp Asp Ser His Gln Ser Asp Glu Ser His His Ser
100 105 110

```
Asp Glu Ser Asp Glu Leu Val Thr Asp Phe Pro Thr Asp Leu Pro Ala
      115                 120                 125

Thr Glu Val Phe Thr Pro Val Val Pro Thr Val Asp Thr Tyr Asp Gly
    130                 135                 140

Arg Gly Asp Ser Val Val Tyr Gly Leu Arg Ser Lys Ser Lys Lys Phe
145                 150                 155                 160

Arg Arg Pro Asp Ile Gln Tyr Pro Asp Ala Thr Asp Glu Asp Ile Thr
                165                 170                 175

Ser His Met Glu Ser Glu Glu Leu Asn Gly Ala Tyr Lys Ala Ile Pro
            180                 185                 190

Val Ala Gln Asp Leu Asn Ala Pro Ser Asp Trp Asp Ser Arg Gly Lys
        195                 200                 205

Asp Ser Tyr Glu Thr Ser Gln Leu Asp Asp Gln Ser Ala Glu Thr His
    210                 215                 220

Ser His Lys Gln Ser Arg Leu Tyr Lys Arg Lys Ala Asn Asp Glu Ser
225                 230                 235                 240

Asn Glu His Ser Asp Val Ile Asp Ser Gln Glu Leu Ser Lys Val Ser
                245                 250                 255

Arg Glu Phe His Ser His Glu Phe His Ser His Glu Asp Met Leu Val
            260                 265                 270

Val Asp Pro Lys Ser Lys Glu Glu Asp Lys His Leu Lys Phe Arg Ile
        275                 280                 285

Ser His Glu Leu Asp Ser Ala Ser Ser Glu Val Asn
    290                 295                 300
```

<210> 49

<211> 99

<212> PRT

<213> Homo sapiens

<400> 49

Met Lys Val Ser Ala Ala Leu Leu Cys Leu Leu Leu Ile Ala Ala Thr
1 5 10 15

Phe Ile Pro Gln Gly Leu Ala Gln Pro Asp Ala Ile Asn Ala Pro Val
20 25 30

Thr Cys Cys Tyr Asn Phe Thr Asn Arg Lys Ile Ser Val Gln Arg Leu
35 40 45

Ala Ser Tyr Arg Arg Ile Thr Ser Ser Lys Cys Pro Lys Glu Ala Val
50 55 60

Ile Phe Lys Thr Ile Val Ala Lys Glu Ile Cys Ala Asp Pro Lys Gln
65 70 75 80

Lys Trp Val Gln Asp Ser Met Asp His Leu Asp Lys Gln Thr Gln Thr
85 90 95

Pro Lys Thr

<210> 50

<211> 470

<212> PRT

<213> Homo sapiens

<400> 50

Met Lys Phe Leu Leu Ile Leu Leu Leu Gln Ala Thr Ala Ser Gly Ala
1 5 10 15

Leu Pro Leu Asn Ser Ser Thr Ser Leu Glu Lys Asn Asn Val Leu Phe
20 25 30

Gly Glu Arg Tyr Leu Glu Lys Phe Tyr Gly Leu Glu Ile Asn Lys Leu
35 40 45

Pro Val Thr Lys Met Lys Tyr Ser Gly Asn Leu Met Lys Glu Lys Ile
50 55 60

```
Gln Glu Met Gln His Phe Leu Gly Leu Lys Val Thr Gly Gln Leu Asp
65                  70                  75                  80

Thr Ser Thr Leu Glu Met Met His Ala Pro Arg Cys Gly Val Pro Asp
                85                  90                  95

Leu His His Phe Arg Glu Met Pro Gly Gly Pro Val Trp Arg Lys His
            100                 105                 110

Tyr Ile Thr Tyr Arg Ile Asn Asn Tyr Thr Pro Asp Met Asn Arg Glu
        115                 120                 125

Asp Val Asp Tyr Ala Ile Arg Lys Ala Phe Gln Val Trp Ser Asn Val
    130                 135                 140

Thr Pro Leu Lys Phe Ser Lys Ile Asn Thr Gly Met Ala Asp Ile Leu
145                 150                 155                 160

Val Val Phe Ala Arg Gly Ala His Gly Asp Phe His Ala Phe Asp Gly
                165                 170                 175

Lys Gly Gly Ile Leu Ala His Ala Phe Gly Pro Gly Ser Gly Ile Gly
            180                 185                 190

Gly Asp Ala His Phe Asp Glu Asp Glu Phe Trp Thr Thr His Ser Gly
        195                 200                 205

Gly Thr Asn Leu Phe Leu Thr Ala Val His Glu Ile Gly His Ser Leu
    210                 215                 220

Gly Leu Gly His Ser Ser Asp Pro Lys Ala Val Met Phe Pro Thr Tyr
225                 230                 235                 240

Lys Tyr Val Asp Ile Asn Thr Phe Arg Leu Ser Ala Asp Asp Ile Arg
                245                 250                 255

Gly Ile Gln Ser Leu Tyr Gly Asp Pro Lys Glu Asn Gln Arg Leu Pro
            260                 265                 270

Asn Pro Asp Asn Ser Glu Pro Ala Leu Cys Asp Pro Asn Leu Ser Phe
        275                 280                 285

Asp Ala Val Thr Thr Val Gly Asn Lys Ile Phe Phe Phe Lys Asp Arg
    290                 295                 300
```

```
Phe Phe Trp Leu Lys Val Ser Glu Arg Pro Lys Thr Ser Val Asn Leu
305                 310                 315                 320

Ile Ser Ser Leu Trp Pro Thr Leu Pro Ser Gly Ile Glu Ala Ala Tyr
                325                 330                 335

Glu Ile Glu Ala Arg Asn Gln Val Phe Leu Phe Lys Asp Asp Lys Tyr
            340                 345                 350

Trp Leu Ile Ser Asn Leu Arg Pro Glu Pro Asn Tyr Pro Lys Ser Ile
        355                 360                 365

His Ser Phe Gly Phe Pro Asn Phe Val Lys Lys Ile Asp Ala Ala Val
    370                 375                 380

Phe Asn Pro Arg Phe Tyr Arg Thr Tyr Phe Phe Val Asp Asn Gln Tyr
385                 390                 395                 400

Trp Arg Tyr Asp Glu Arg Arg Gln Met Met Asp Pro Gly Tyr Pro Lys
                405                 410                 415

Leu Ile Thr Lys Asn Phe Gln Gly Ile Gly Pro Lys Ile Asp Ala Val
            420                 425                 430

Phe Tyr Ser Lys Asn Lys Tyr Tyr Tyr Phe Phe Gln Gly Ser Asn Gln
        435                 440                 445

Phe Glu Tyr Asp Phe Leu Leu Gln Arg Ile Thr Lys Thr Leu Lys Ser
    450                 455                 460

Asn Ser Trp Phe Gly Cys
465                 470
```

<210> 51

<211> 101

<212> PRT

<213> Homo sapiens

<400> 51

Met Gly Thr Arg Leu Leu Pro Ala Leu Phe Leu Val Leu Leu Val Leu
1 5 10 15

Gly Phe Glu Val Gln Gly Thr Gln Gln Pro Gln Gln Asp Glu Met Pro
20 25 30

Ser Pro Thr Phe Leu Thr Gln Val Lys Glu Ser Leu Ser Ser Tyr Trp
35 40 45

Glu Ser Ala Lys Thr Ala Ala Gln Asn Leu Tyr Glu Lys Thr Tyr Leu
50 55 60

Pro Ala Val Asp Glu Lys Leu Arg Asp Leu Tyr Ser Lys Ser Thr Ala
65 70 75 80

Ala Met Ser Thr Tyr Thr Gly Ile Phe Thr Asp Gln Val Leu Ser Val
85 90 95

Leu Lys Gly Glu Glu
100

<210> 52

<211> 150

<212> PRT

<213> Homo sapiens

<400> 52

Met Phe Arg Lys Gly Lys Lys Arg His Ser Ser Ser Ser Ser Gln Ser
1 5 10 15

Ser Glu Ile Ser Thr Lys Ser Lys Ser Val Asp Ser Ser Leu Gly Gly
20 25 30

Leu Ser Arg Ser Ser Thr Val Ala Ser Leu Asp Thr Asp Ser Thr Lys
35 40 45

Ser Ser Gly Gln Ser Asn Asn Asn Ser Asp Thr Cys Ala Glu Phe Arg
50 55 60

Ile Lys Tyr Val Gly Ala Ile Glu Lys Leu Lys Leu Ser Glu Gly Lys
65 70 75 80

Gly Leu Glu Gly Pro Leu Asp Leu Ile Asn Tyr Ile Asp Val Ala Gln
85 90 95

Gln Asp Gly Lys Leu Pro Phe Val Pro Pro Glu Glu Glu Phe Ile Met
100 105 110

Gly Val Ser Lys Tyr Gly Ile Lys Val Ser Thr Ser Asp Gln Tyr Glu
115 120 125

Gln Ala Gln Ala Ile Cys Lys Val Leu Ser Thr Ala Phe Asp Ser Val
130 135 140

Leu Thr Ser Glu Lys Pro
145 150

<210> 53

<211> 381

<212> PRT

<213> Homo sapiens

<400> 53

Met Ile Asp Thr Leu Arg Pro Val Pro Phe Ala Ser Glu Met Ala Ile
1 5 10 15

Ser Lys Thr Val Ala Trp Leu Asn Glu Gln Leu Glu Leu Gly Asn Glu
20 25 30

Arg Leu Leu Leu Met Asp Cys Arg Pro Gln Glu Leu Tyr Glu Ser Ser
35 40 45

His Ile Glu Ser Ala Ile Asn Val Ala Ile Pro Gly Ile Met Leu Arg
50 55 60

Arg Leu Gln Lys Gly Asn Leu Pro Val Arg Ala Leu Phe Thr Arg Gly
65 70 75 80

Glu Asp Arg Asp Arg Phe Thr Arg Arg Cys Gly Thr Asp Thr Val Val
85 90 95

Leu Tyr Asp Glu Ser Ser Ser Asp Trp Asn Glu Asn Thr Gly Gly Glu
100 105 110

Ser Leu Leu Gly Leu Leu Leu Lys Lys Leu Lys Asp Glu Gly Cys Arg
115 120 125

Ala Phe Tyr Leu Glu Gly Gly Phe Ser Lys Phe Gln Ala Glu Phe Ser
130 135 140

Leu His Cys Glu Thr Asn Leu Asp Gly Ser Cys Ser Ser Ser Ser Pro
145 150 155 160

Pro Leu Pro Val Leu Gly Leu Gly Gly Leu Arg Ile Ser Ser Asp Ser
165 170 175

Ser Ser Asp Ile Glu Ser Asp Leu Asp Arg Asp Pro Asn Ser Ala Thr
180 185 190

Asp Ser Asp Gly Ser Pro Leu Ser Asn Ser Gln Pro Ser Phe Pro Val
195 200 205

Glu Ile Leu Pro Phe Leu Tyr Leu Gly Cys Ala Lys Asp Ser Thr Asn
210 215 220

Leu Asp Val Leu Glu Glu Phe Gly Ile Lys Tyr Ile Leu Asn Val Thr
225 230 235 240

Pro Asn Leu Pro Asn Leu Phe Glu Asn Ala Gly Glu Phe Lys Tyr Lys
245 250 255

Gln Ile Pro Ile Ser Asp His Trp Ser Gln Asn Leu Ser Gln Phe Phe
260 265 270

Pro Glu Ala Ile Ser Phe Ile Asp Glu Ala Arg Gly Lys Asn Cys Gly
275 280 285

Val Leu Val His Cys Leu Ala Gly Ile Ser Arg Ser Val Thr Val Thr
290 295 300

Val Ala Tyr Leu Met Gln Lys Leu Asn Leu Ser Met Asn Asp Ala Tyr
305 310 315 320

Asp Ile Val Lys Met Lys Lys Ser Asn Ile Ser Pro Asn Phe Asn Phe
325 330 335

Met Gly Gln Leu Leu Asp Phe Glu Arg Thr Leu Gly Leu Ser Ser Pro
340 345 350

Cys Asp Asn Arg Val Pro Ala Gln Gln Leu Tyr Phe Thr Thr Pro Ser
355 360 365

Asn Gln Asn Val Tyr Gln Val Asp Ser Leu Gln Ser Thr
370 375 380

<210> 54

<211> 542

<212> PRT

<213> Homo sapiens

<400> 54

Met Thr Lys Ser Asn Gly Glu Glu Pro Lys Met Gly Gly Arg Met Glu
1 5 10 15

Arg Phe Gln Gln Gly Val Arg Lys Arg Thr Leu Leu Ala Lys Lys Lys
20 25 30

Val Gln Asn Ile Thr Lys Glu Asp Val Lys Ser Tyr Leu Phe Arg Asn
35 40 45

Ala Phe Val Leu Leu Thr Val Thr Ala Val Ile Val Gly Thr Ile Leu
50 55 60

Gly Phe Thr Leu Arg Pro Tyr Arg Met Ser Tyr Arg Glu Val Lys Tyr
65 70 75 80

Phe Ser Phe Pro Gly Glu Leu Leu Met Arg Met Leu Gln Met Leu Val
85 90 95

Leu Pro Leu Ile Ile Ser Ser Leu Val Thr Gly Met Ala Ala Leu Asp
100 105 110

Ser Lys Ala Ser Gly Lys Met Gly Met Arg Ala Val Val Tyr Tyr Met
115 120 125

```
Thr Thr Thr Ile Ile Ala Val Val Ile Gly Ile Ile Ile Val Ile Ile
    130                 135                 140

Ile His Pro Gly Lys Gly Thr Lys Glu Asn Met His Arg Glu Gly Lys
145                 150                 155                 160

Ile Val Arg Val Thr Ala Ala Asp Ala Phe Leu Asp Leu Ile Arg Asn
                165                 170                 175

Met Phe Pro Pro Asn Leu Val Glu Ala Cys Phe Lys Gln Phe Lys Thr
            180                 185                 190

Asn Tyr Glu Lys Arg Ser Phe Lys Val Pro Ile Gln Ala Asn Glu Thr
        195                 200                 205

Leu Val Gly Ala Val Ile Asn Asn Val Ser Glu Ala Met Glu Thr Leu
    210                 215                 220

Thr Arg Ile Thr Glu Glu Leu Val Pro Val Pro Gly Ser Val Asn Gly
225                 230                 235                 240

Val Asn Ala Leu Gly Leu Val Val Phe Ser Met Cys Phe Gly Phe Val
                245                 250                 255

Ile Gly Asn Met Lys Glu Gln Gly Gln Ala Leu Arg Glu Phe Phe Asp
            260                 265                 270

Ser Leu Asn Glu Ala Ile Met Arg Leu Val Ala Val Ile Met Trp Tyr
        275                 280                 285

Ala Pro Val Gly Ile Leu Phe Leu Ile Ala Gly Lys Ile Val Glu Met
    290                 295                 300

Glu Asp Met Gly Val Ile Gly Gly Gln Leu Ala Met Tyr Thr Val Thr
305                 310                 315                 320

Val Ile Val Gly Leu Leu Ile His Ala Val Ile Val Leu Pro Leu Leu
                325                 330                 335

Tyr Phe Leu Val Thr Arg Lys Asn Pro Trp Val Phe Ile Gly Gly Leu
            340                 345                 350

Leu Gln Ala Leu Ile Thr Ala Leu Gly Thr Ser Ser Ser Ser Ala Thr
        355                 360                 365
```

```
Leu Pro Ile Thr Phe Lys Cys Leu Glu Glu Asn Asn Gly Val Asp Lys
    370                 375                 380

Arg Val Thr Arg Phe Val Leu Pro Val Gly Ala Thr Ile Asn Met Asp
385                 390                 395                 400

Gly Thr Ala Leu Tyr Glu Ala Leu Ala Ala Ile Phe Ile Ala Gln Val
                405                 410                 415

Asn Asn Phe Glu Leu Asn Phe Gly Gln Ile Ile Thr Ile Ser Ile Thr
            420                 425                 430

Ala Thr Ala Ala Ser Ile Gly Ala Ala Gly Ile Pro Gln Ala Gly Leu
        435                 440                 445

Val Thr Met Val Ile Val Leu Thr Ser Val Gly Leu Pro Thr Asp Asp
    450                 455                 460

Ile Thr Leu Ile Ile Ala Val Asp Trp Phe Leu Asp Arg Leu Arg Thr
465                 470                 475                 480

Thr Thr Asn Val Leu Gly Asp Ser Leu Gly Ala Gly Ile Val Glu His
                485                 490                 495

Leu Ser Arg His Glu Leu Lys Asn Arg Asp Val Glu Met Gly Asn Ser
            500                 505                 510

Val Ile Glu Glu Asn Glu Met Lys Lys Pro Tyr Gln Leu Ile Ala Gln
        515                 520                 525

Asp Asn Glu Thr Glu Lys Pro Ile Asp Ser Glu Thr Lys Met
    530                 535                 540
```

<210> 55

<211> 328

<212> PRT

<213> Homo sapiens

<400> 55

```
Met Leu Pro Arg Val Gly Cys Pro Ala Leu Pro Leu Pro Pro Pro Pro
1               5                   10                  15
```

```
Leu Leu Pro Leu Leu Pro Leu Leu Leu Leu Leu Leu Gly Ala Ser Gly
            20                  25                  30

Gly Gly Gly Gly Ala Arg Ala Glu Val Leu Phe Arg Cys Pro Pro Cys
        35                  40                  45

Thr Pro Glu Arg Leu Ala Ala Cys Gly Pro Pro Pro Val Ala Pro Pro
    50                  55                  60

Ala Ala Val Ala Ala Val Ala Gly Gly Ala Arg Met Pro Cys Ala Glu
65                  70                  75                  80

Leu Val Arg Glu Pro Gly Cys Gly Cys Cys Ser Val Cys Ala Arg Leu
                85                  90                  95

Glu Gly Glu Ala Cys Gly Val Tyr Thr Pro Arg Cys Gly Gln Gly Leu
            100                 105                 110

Arg Cys Tyr Pro His Pro Gly Ser Glu Leu Pro Leu Gln Ala Leu Val
        115                 120                 125

Met Gly Glu Gly Thr Cys Glu Lys Arg Arg Asp Ala Glu Tyr Gly Ala
    130                 135                 140

Ser Pro Glu Gln Val Ala Asp Asn Gly Asp Asp His Ser Glu Gly Gly
145                 150                 155                 160

Leu Val Glu Asn His Val Asp Ser Thr Met Asn Met Leu Gly Gly Gly
                165                 170                 175

Gly Ser Ala Gly Arg Lys Pro Leu Lys Ser Gly Met Lys Glu Leu Ala
            180                 185                 190

Val Phe Arg Glu Lys Val Thr Glu Gln His Arg Gln Met Gly Lys Gly
        195                 200                 205

Gly Lys His His Leu Gly Leu Glu Glu Pro Lys Lys Leu Arg Pro Pro
    210                 215                 220

Pro Ala Arg Thr Pro Cys Gln Gln Glu Leu Asp Gln Val Leu Glu Arg
225                 230                 235                 240
```

```
Ile Ser Thr Met Arg Leu Pro Asp Glu Arg Gly Pro Leu Glu His Leu
                245                 250                 255

Tyr Ser Leu His Ile Pro Asn Cys Asp Lys His Gly Leu Tyr Asn Leu
            260                 265                 270

Lys Gln Cys Lys Met Ser Leu Asn Gly Gln Arg Gly Glu Cys Trp Cys
        275                 280                 285

Val Asn Pro Asn Thr Gly Lys Leu Ile Gln Gly Ala Pro Thr Ile Arg
    290                 295                 300

Gly Asp Pro Glu Cys His Leu Phe Tyr Asn Glu Gln Gln Glu Ala Cys
305                 310                 315                 320

Gly Val His Thr Gln Arg Met Gln
                325
```

<210> 56

<211> 38

<212> PRT

<213> Homo sapiens

<400> 56

```
Met Gln Ser Met Pro Gln Ser Pro Ala Val Ile Thr Ser Pro Ile Gly
1               5                   10                  15

Arg Ser Gln Cys Arg Gly Ser Arg Ala Ile Glu Glu Ser Pro Ala Ala
            20                  25                  30

Ser Val Pro Asn Lys Leu
        35
```

<210> 57

<211> 29

<212> DNA

<213> Artificial DNA

<400> 57
acgttcagac cacctattcc cttcttgcg 29

<210> 58

<211> 24

<212> DNA

<213> Artificial DNA

<400> 58
attcctgctt ttcagagtga gaca 24

<210> 59

<211> 24

<212> DNA

<213> Artificial DNA

<400> 59
gaacagcttc cctcactgtg taca 24

<210> 60

<211> 19

<212> DNA

<213> Artificial DNA

<400> 60
tcgccaggcc ctgtggtgg 19

<210> 61

<211> 16

<212> DNA

<213> Artificial DNA

<400> 61
agggcgttcc gcaggt 16

<210> 62

<211> 23

<212> DNA

<213> Artificial DNA

<400> 62
gggcctttct gtgaaagttg taa 23

<210> 63

<211> 23

<212> DNA

<213> Artificial DNA

<400> 63
ttcccagcat catccaggcc cag 23

<210> 64

<211> 21

<212> DNA

<213> Artificial DNA

<400> 64
cgagcccttt gatgacttcc t 21

<210> 65

<211> 20

<212> DNA

<213> Artificial DNA

<400> 65
ggctcccagt ctgctgcata 20

<210> 66

<211> 24

<212> DNA

<213> Artificial DNA

<400> 66
cgcgtctctc aagctcgcct cttc 24

<210> 67

<211> 22

<212> DNA

<213> Artificial DNA

<400> 67
cagctacttt tctggtcagg gc 22

<210> 68

<211> 20

<212> DNA

<213> Artificial DNA

<400> 68
acaatcaggc gttccagctc 20

<210> 69

<211> 31

<212> DNA

<213> Artificial DNA

<400> 69
acagtggagc tctgtattag aaagcccctc a 31

<210> 70

<211> 29

<212> DNA

<213> Artificial DNA

<400> 70
gaggaagata ctgtggtact gtcataaaa 29

<210> 71

<211> 22

<212> DNA

<213> Artificial DNA

<400> 71
gagttacctg gccttcccag tt 22

<210> 72

<211> 25

<212> DNA

<213> Artificial DNA

<400> 72
ccaacttcag cttcatgggc cagct 25

<210> 73

<211> 20

<212> DNA

<213> Artificial DNA

<400> 73
gcagaggcga agcatcatct 20

<210> 74

<211> 21

<212> DNA

<213> Artificial DNA

<400> 74
cagcacctgg gactcaaact g 21

<210> 75

<211> 22

<212> DNA

<213> Artificial DNA

<400> 75
tgccagtgct tgcagaccct gc 22

<210> 76

<211> 18

<212> DNA

<213> Artificial DNA

<400> 76
cctggccact gaactgcg 18

<210> 77

<211> 23

<212> DNA

<213> Artificial DNA

<400> 77
tggatgttct tgaggtgaat tcc 23

<210> 78

<211> 26

<212> DNA

<213> Artificial DNA

<400> 78
tgaattccct gcagtgtctg caagca 26

<210> 79

<211> 25

<212> DNA

<213> Artificial DNA

<400> 79
acattcacac tttggatgtt cttga 25

<210> 80

<211> 19

<212> DNA

<213> Artificial DNA

<400> 80
gagcgtccgt ggtcactga 19

<210> 81

<211> 35

<212> DNA

<213> Artificial DNA

<400> 81
ttctgactct aagtggcatt caaggagtac ctctc 35

<210> 82

<211> 23

<212> DNA

<213> Artificial DNA

<400> 82
cgattctgat ttgctgcctt atc 23

<210> 83

<211> 21

<212> DNA

<213> Artificial DNA

<400> 83
gcaggtacag cgtacggttc t 21

<210> 84

<211> 25

<212> DNA

<213> Artificial DNA

<400> 84
tgatgcggtc tgtccattgc cgata 25

<210> 85

<211> 22

<212> DNA

<213> Artificial DNA

<400> 85
ttgtctccct gctggaaaaa tt 22

<210> 86

<211> 19

<212> DNA

<213> Artificial DNA

<400> 86
tgagcaaccc caccaagtc 19

<210> 87

<211> 22

<212> DNA

<213> Artificial DNA

<400> 87
ccgcatccag aagttccgcg tc 22

<210> 88

<211> 18

<212> DNA

<213> Artificial DNA

<400> 88
gcaactcggc catgaacc 18

<210> 89

<211> 24

<212> DNA

<213> Artificial DNA

<400> 89
ggtcattcca aatcttaagg aagg 24

<210> 90

<211> 23

<212> DNA

<213> Artificial DNA

<400> 90
ccggtccttc taccccgagc aca 23

<210> 91

<211> 22

<212> DNA

<213> Artificial DNA

<400> 91
ctgcgtccaa caatgagacc ta 22

<210> 92

<211> 20

<212> DNA

<213> Artificial DNA

<400> 92
ccgatcagcc actccttgat 20

<210> 93

<211> 29

<212> DNA

<213> Artificial DNA

<400> 93
aggaccttat ctgtgcgtga aacacacca 29

<210> 94

<211> 25

<212> DNA

<213> Artificial DNA

<400> 94
cctagagggt tcctgtagac ctagg 25

<210> 95

<211> 19

<212> DNA

<213> Artificial DNA

<400> 95
agtccttgag gcccacagc 19

<210> 96

<211> 22

<212> DNA

<213> Artificial DNA

<400> 96
cgccgccttg gcatagtttg ga 22

<210> 97

<211> 21

<212> DNA

<213> Artificial DNA

<400> 97
tccaaggaga ccagaagacc a 21

<210> 98

<211> 20

<212> DNA

<213> Artificial DNA

<400> 98
gtgtggaagg agtggctgct 20

<210> 99

<211> 27

<212> DNA

<213> Artificial DNA

<400> 99
tagatctgga aaggacggcg tcatgga 27

<210> 100

<211> 16

<212> DNA

<213> Artificial DNA

<400> 100
gcaggcacac ggtgca 16

<210> 101

<211> 20

<212> DNA

<213> Artificial DNA

<400> 101
gcaggctcga tgctgttcat 20

<210> 102

<211> 29

<212> DNA

<213> Artificial DNA

<400> 102
ctccaaagga aagcgaggtg gacattcat 29

<210> 103

<211> 19

<212> DNA

<213> Artificial DNA

<400> 103
gggtccagcc tcaagatcc 19

<210> 104

<211> 26

<212> DNA

<213> Artificial DNA

<400> 104
acttgaagac tcagactcag tggaaa 26

<210> 105

<211> 26

<212> DNA

<213> Artificial DNA

<400> 105
cgtgtcctct ccctgccaag tctcag 26

<210> 106

<211> 29

<212> DNA

<213> Artificial DNA

<400> 106
gtcaaagtgt aacagatatc agtgtctcc 29

<210> 107

<211> 20

<212> DNA

<213> Artificial DNA

<400> 107
tacaggcatg gaagcccaac 20

<210> 108

<211> 31

<212> DNA

<213> Artificial DNA

<400> 108
cctacaatga tgcagcactt ctcacagttc c 31

<210> 109

<211> 24

<212> DNA

<213> Artificial DNA

<400> 109
gaagccagaa tacttggagg aact 24

<210> 110

<211> 22

<212> DNA

<213> Artificial DNA

<400> 110
atacggtgga ggtcaaggac at 22

<210> 111

<211> 25

<212> DNA

<213> Artificial DNA

<400> 111
attcctcatg ccaccactgc caaca 25

<210> 112

<211> 20

<212> DNA

<213> Artificial DNA

<400> 112
tccagctttg tgcctgtcac 20

<210> 113

<211> 20

<212> DNA

<213> Artificial DNA

<400> 113
gggaatgtgg tagcccaaga 20

<210> 114

<211> 27

<212> DNA

<213> Artificial DNA

<400> 114
tcacctcaca catggaaagc gaggagt 27

<210> 115

<211> 23

<212> DNA

<213> Artificial DNA

<400> 115
ccagtaccct gatgctacag acg 23

<210> 116

<211> 20

<212> DNA

<213> Artificial DNA

<400> 116
tggccttgta tgcaccattc 20

<210> 117

<211> 29

<212> DNA

<213> Artificial DNA

<400> 117
acccaaactc cgaagacttg aacacccac 29

<210> 118

<211> 19

<212> DNA

<213> Artificial DNA

<400> 118
tggaccacct ggacaagca 19

<210> 119

<211> 21

<212> DNA

<213> Artificial DNA

<400> 119
gctgcagatt cttgggttgt g 21

<210> 120

<211> 25

<212> DNA

<213> Artificial DNA

<400> 120
ccttatggcc aaccttgcca tctgg 25

<210> 121

<211> 26

<212> DNA

<213> Artificial DNA

<400> 121
ctgagagacc aaagaccagt gttaat 26

<210> 122

<211> 24

<212> DNA

<213> Artificial DNA

<400> 122
tttctggctt caatttcata agca 24

<210> 123

<211> 25

<212> DNA

<213> Artificial DNA

<400> 123
acagccgccc agaacctgta cgaga 25

<210> 124

<211> 22

<212> DNA

<213> Artificial DNA

<400> 124
tccagttact gggagtcagc aa 22

<210> 125

<211> 21

<212> DNA

<213> Artificial DNA

<400> 125
tctacagcgg gcaggtatgt c 21

<210> 126

<211> 23

<212> DNA

<213> Artificial DNA

<400> 126
ttcacgatcc agcactgtgg cca 23

<210> 127

<211> 24

<212> DNA

<213> Artificial DNA

<400> 127
agcaagtctg tggattctag cctt 24

<210> 128

<211> 22

<212> DNA

<213> Artificial DNA

<400> 128
cctgagcttt tggtggaatc tg 22

<210> 129

<211> 36

<212> DNA

<213> Artificial DNA

<400> 129
atgcttgact ttaccaattc tgatgacatc tttacg 36

<210> 130

<211> 25

<212> DNA

<213> Artificial DNA

<400> 130
gaaaattgtg ctctgttgta atcca 25

<210> 131

<211> 29

<212> DNA

<213> Artificial DNA

<400> 131
gccataacaa ggtcttagtg ataatagtg 29

<210> 132

<211> 24

<212> DNA

<213> Artificial DNA

<400> 132
tggactggtt cctggatcgc ctcc 24

<210> 133

<211> 21

<212> DNA

<213> Artificial DNA

<400> 133
cactgacgac atcacgctca t 21

<210> 134

<211> 20

<212> DNA

<213> Artificial DNA

<400> 134
cccagtacgt tggtggtggt 20

<210> 135

<211> 31

<212> DNA

<213> Artificial DNA

<400> 135
aacctcaaac agtgcaagat gtctctgaac g 31

<210> 136

<211> 21

<212> DNA

<213> Artificial DNA

<400> 136
ctgtgacaag catggcctgt a 21

<210> 137

<211> 16

<212> DNA

<213> Artificial DNA

<400> 137
gcactcccca cgctgc 16

<210> 138

<211> 21

<212> DNA

<213> Artificial DNA

<400> 138
ccatggcagc cctttggtgc a 21

<210> 139

<211> 21

<212> DNA

<213> Artificial DNA

<400> 139
acctgcgttc ctcctctagc t 21

<210> 140

<211> 22

<212> DNA

<213> Artificial DNA

<400> 140
tggtctagaa gtgcagccca tt 22

<210> 141

<211> 24

<212> DNA

<213> Artificial DNA

<400> 141
gaacagcttc cctcactgtg taca 24

<210> 142

<211> 25

<212> DNA

<213> Artificial DNA

<400> 142
attcctgctt ttcagagtga gacac 25

<210> 143

<211> 29

<212> DNA

<213> Artificial DNA

<400> 143
acgttcagac cacctattcc cttcttgcg 29

<210> 144

<211> 21

<212> DNA

<213> Artificial DNA

<400> 144
tctttggcta ccagtccagc a 21

<210> 145

<211> 20

<212> DNA

<213> Artificial DNA

<400> 145
aaaggcccga gtatccctga 20

<210> 146

<211> 29

<212> DNA

<213> Artificial DNA

<400> 146
caagtcaccc agtcctactg ggacaccaa 29

<210> 147

<211> 22

<212> DNA

<213> Artificial DNA

<400> 147
tctaggactt ctgcacggac ct 22

<210> 148

<211> 20

<212> DNA

<213> Artificial DNA

<400> 148
aggtccggac tggtcgagat 20

<210> 149

<211> 27

<212> DNA

<213> Artificial DNA

<400> 149
cagtgccaac ttcattccca cggtcac 27

<210> 150

<211> 22

<212> DNA

<213> Artificial DNA

<400> 150
cagctacttt tctggtcagg gc 22

<210> 151

<211> 20

<212> DNA

<213> Artificial DNA

<400> 151
acaatcaggc gttccagctc 20

<210> 152

<211> 24

<212> DNA

<213> Artificial DNA

<400> 152
cgcgtctctc aagctcgcct cttc 24

<210> 153

<211> 25

<212> DNA

<213> Artificial DNA

<400> 153
tggctgaagg gaattaacag atagt 25

<210> 154

<211> 27

<212> DNA

<213> Artificial DNA

<400> 154
ttcatgttaa tccattctga gtcattt 27

<210> 155

<211> 26

<212> DNA

<213> Artificial DNA

<400> 155
tcacccaggg agccatactg gcacat 26

<210> 156

<211> 20

<212> DNA

<213> Artificial DNA

<400> 156
gcagaggcga agcatcatct 20

<210> 157

<211> 21

<212> DNA

<213> Artificial DNA

<400> 157
cagcacctgg gactcaaact g 21

<210> 158

<211> 25

<212> DNA

<213> Artificial DNA

<400> 158
ccaacttcag cttcatgggc cagct 25

<210> 159

<211> 18

<212> DNA

<213> Artificial DNA

<400> 159
tgctgctcct gctcctgg 18

<210> 160

<211> 20

<212> DNA

<213> Artificial DNA

<400> 160
aggtgaattc cctgcagggt 20

<210> 161

<211> 25

<212> DNA

<213> Artificial DNA

<400> 161
cactgaactg cgctgccagt gcttg 25

<210> 162

<211> 19

<212> DNA

<213> Artificial DNA

<400> 162
gagcgtccgt ggtcactga 19

<210> 163

<211> 25

<212> DNA

<213> Artificial DNA

<400> 163
acattcacac tttggatgtt cttga 25

<210> 164

<211> 26

<212> DNA

<213> Artificial DNA

<400> 164
tgaattccct gcagtgtctg caagca 26

<210> 165

<211> 20

<212> DNA

<213> Artificial DNA

<400> 165
tgccagctat gtggtagggc 20

<210> 166

<211> 20

<212> DNA

<213> Artificial DNA

<400> 166
gtcaaccgtg ggatttcctg 20

<210> 167

<211> 31

<212> DNA

<213> Artificial DNA

<400> 167
cctggccttt cagttacttg agtttggagt c 31

<210> 168

<211> 22

<212> DNA

<213> Artificial DNA

<400> 168
cgtgcaaagt cacacacctc at 22

<210> 169

<211> 23

<212> DNA

<213> Artificial DNA

<400> 169
ggtgttggtg aagtaggtga tgg 23

<210> 170

<211> 25

<212> DNA

<213> Artificial DNA

<400> 170
cttcctcacg tgcatgagcg tggac 25

<210> 171

<211> 19

<212> DNA

<213> Artificial DNA

<400> 171
cctttgaagg gcaggcaca 19

<210> 172

<211> 20

<212> DNA

<213> Artificial DNA

<400> 172
gcaggctcga tgctgttcat 20

<210> 173

<211> 27

<212> DNA

<213> Artificial DNA

<400> 173
tccatgacgc cgtcctttcc agatcta 27

<210> 174

<211> 24

<212> DNA

<213> Artificial DNA

<400> 174
ctgagaagca tgacggacaa gtac 24

<210> 175

<211> 22

<212> DNA

<213> Artificial DNA

<400> 175
agggaagcgt gatgacaaag ag 22

<210> 176

<211> 24

<212> DNA

<213> Artificial DNA

<400> 176
ctgcacctgt cagtggccga cctc 24

<210> 177

<211> 22

<212> DNA

<213> Artificial DNA

<400> 177
aggaacaggg aacattggag ct 22

<210> 178

<211> 19

<212> DNA

<213> Artificial DNA

<400> 178
ttggtgaggt gggtgggat 19

<210> 179

<211> 33

<212> DNA

<213> Artificial DNA

<400> 179
aactgattac ctgccaattg ctaccgagaa ggt 33

<210> 180

<211> 20

<212> DNA

<213> Artificial DNA

<400> 180
agaggaggtc gcagttcagc 20

<210> 181

<211> 22

<212> DNA

<213> Artificial DNA

<400> 181
cagtacccca gtgtcatgga ca 22

<210> 182

<211> 27

<212> DNA

<213> Artificial DNA

<400> 182
tctgcagaat cgacaccaac cagcatc 27

CLAIMS

1. A method for the prediction, diagnosis or prognosis of chronic lung disease by the detection of:

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

in a biological sample comprising the following steps:

hybridizing at least one polynucleotide specified in (a) to (d) to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and

detecting said hybridization complex.

2. The method of claim 1, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

3. A method for the prediction, diagnosis or prognosis of chronic lung disease by the detection of:

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

e) a purified polypeptide encoded by a polynucleotide sequence specified in (a) to (d);

f) a purified polypeptide comprising at least one of the sequences of SEQ ID NO. 29 to 56;

comprising the steps of contacting a biological sample with a reagent which specifically interacts with the polynucleotide specified in (a) to (d) or the polypeptide specified in (e) and (f).

4. A diagnostic kit for conducting the method of anyone of claims 1 to 3.

5. A method of screening for agents which regulate the activity of a polypeptide encoded by a polynucleotide selected from the group consisting of:

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

comprising the steps of:

contacting a test compound with at least one polypeptide encoded by a polynucleotide specified in (a) to (d); and

detecting binding of the test compound to the polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for modulating the activity of the polypeptide in order to prevent or treat chronic lung disease.

6. A method of screening for agents which regulate the activity of a polypeptide encoded by a polynucleotide selected from the group consisting of:

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

comprising the steps of:

contacting a test compound with at least one polypeptide encoded by a polynucleotide specified in (a) to (d); and

detecting the activity of the polypeptide as specified for the respective sequence in Table 3, wherein a test compound which increases the activity is identified as a potential preventive or therapeutic agent for increasing the in chronic lung disease, and wherein a test compound which decreases the activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity in chronic lung disease.

7. A method of screening for agents which regulate the activity of a polynucleotide selected from group consisting of;

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

comprising the steps of:

contacting a test compound with at least one polynucleotide specified in (a) to (d), and

detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential preventive or therapeutic agent for regulating the activity of the polynucleotide in chronic lung disease.

8. Use of

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and encodes a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

e) an antisense molecule targeting one of the polynucleotide sequences specified in (a) to (d);

f) a purified polypeptide encoded by a polynucleotide sequence specified in (a) to (d)

g) a purified polypeptide comprising at least one of the sequences of SEQ ID NO. 29 to 56;

h) an antibody capable of binding to one of the polynucleotide specified in (a) to (d) or a polypeptide specified in (f) and (g)

i) a reagent identified by any of the methods of claim 5 to 7 that modulates the amount or activity of a polynucleotide sequence specified in (a) to (d) or a polypeptide specified in (f) and (g)

for the prevention, prediction, diagnosis, prognosis and treatment of chronic lung disease.

9. Use of claim 8 wherein the disease is COPD.

10. A reagent that regulates the activity of a polynucleotide selected from the group consisting of:

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

or a polypeptide encoded by at least one of the polynucleotides specified in (a) to (d). wherein said reagent is identified by the method of any of the claims 5 to 7.

11. A pharmaceutical composition, comprising:

an expression vector containing at least one polynucleotide selected from the group consisting of:

a) a polynucleotide comprising at least one of the sequences of SEQ ID NO. 1 to 28;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 3;

c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the degeneracy of the genetic code;

d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

or the reagent of claim 10 and a pharmaceutically acceptable carrier.

12. A method of treating or preventing chronic lung disease by administering to a mammal in need thereof an agent identified by the method of claim 5, 6 or 7.

Table 1: Genes, up-regulated in COPD patients

No	Accession-No	Fold Change		Description
		Arrays	TaqMan	
1	L 13463	25,2	21.0	Cluster Incl L13463:Human helix-loop-helix basic phosphoprotein (G0S8) mRNA, complete cds
2	X 14787	15,5	16,4	X14787 //DEFINITION=HSTS Human mRNA for thrombospondin
3	K 00650	15,1	5,4	K00650 //DEFINITION=HUMFOS Human fos proto-oncogene (c-fos), complete cds
4	M 29039	6,2	4,3	M29039 //DEFINITION=HUMJUNCAA Human transactivator (jun-B) gene, complete cds
5	S 59049	5,9	12,6	Cluster Incl S59049:BL34=B cell activation gene [human, mRNA, 1398 nt]
6	X 68277	4,4	5,2	X68277//HSCL100 H.sapiens CL 100 mRNA for protein tyrosine phosphatase
7	M 36820	4,3	5,5	Cluster Incl M36820:Human cytokine (GRO-beta) mRNA, complete cds
8	M 36821	3,8	5,3	Cluster Incl M36821:Human cytokine (GRO-gamma) mRNA, complete cds
9	X02530	3,5	7,8	X02530 /Human mRNA for gamma-interferon inducible early response gene (with homology to platelet proteins)
10	L 20971	3,3	4,1	Cluster Incl L20971:Human phosphodiesterase mRNA, complete cds
11	L 22214	3,3	3,4	Cluster Incl L22214:Human adenosine A1 receptor (ADORA1) mRNA exons 1-6, complete cds
12	U 67784	3,1	4,4	Cluster Incl U67784:Human orphan G protein-coupled receptor (RDC1) mRNA, partial cds
13	V 01512	3,0	5,9	V01512 //DEFINITION=HSCFOS Human cellular oncogene c-fos (complete sequence)
14	AL 021154	2,6	5,1	Cluster Incl AL021154:(Inhibitor of DNA binding 3 (dominant negative helix-loop-helix protein, 1R21, HEIR-1))
15	M 20681	2,4	2,1	Cluster Incl M20681:Human glucose transporter-like protein-III (GLUT3), complete cds
16	L 06797	2,4	2,9	L06797/HUMGPCR Human (clone L5) orphan G protein-coupled receptor mRNA, complete cds
17	M 60278	2,1	2,2	Cluster Incl M60278:Human heparin-binding EGF-like growth factor mRNA, complete cds
18	M 96233	2,0	1,4	M96233 /Human glutathione transferase class mu number 4 (GSTM4) gene, complete cds

Table 2: Genes, down-regulated in COPD patients

No	Accession-No	Fold Change		Description
		Arrays	TaqMan	
19	U03688	-8.7	-8.11	Cluster Incl U03688:Human dioxin-inducible cytochrome P450 (CYP1B1) mRNA, complete cds
20	AF 052124	-5.9	-5,2	Cluster Incl AF052124:Homo sapiens osteopontin mRNA, complete cds
21	M 28225	-3.8	-6,9	Cluster Incl M28225:Human JE gene encoding a monocyte secretory protein
22	L 23808	-3.2	-4,4	L23808 /DEFINITION=HUMHME Human metalloproteinase (HME) mRNA, complete cds
23	AA 883870	-2.4	-1,8	Cluster Incl AA883870, similar to gb:M29844 APOLIPOPROTEIN C-II PRECURSOR (HUMAN)
24	AF 012024	-2.1	-1.0	AF012024/ integrin cytoplasmic domain associated protein (Icap-1b) mRNA, complete cds
25	AB 13382	-2.0	-2,0	Cluster Incl AB013382:Homo sapiens mRNA for DUSP6, complete cds
26	D 26443	-1.98	-2,0	Cluster Incl D26443:Human mRNA for glutamate transporter, complete cds
27	X16302	-1.51	-1.24	Cluster Incl X16302:Human mRNA for insulin-like growth factor binding protein (IGFBP-2)
28	M26683	-1.51	-10.13	M26683 / Human interferon gamma treatment inducible mRNA

Table 3: Biochemical function of the „COPD GENE"

No	mRNA Accession-No	Protein Accession-No	Biochemical Function of the „COPD GENE"
1	L 13463	AAC37587	GTPase activating protein, Regulator of G-protein signaling 2, one of a family of proteins that negatively regulates G protein-coupled receptor signaling by binding to G-protein alpha subunits and stimulating their intrinsic GTPase activity
2	X 14787	CAA32889	Inhibitor or repressor; Adhesin/agglutinin, Thrombospondin-1, a member of a family of adhesive molecules, expressed in platelets and in extracellular matrix of developing and repairing tissues, has a role in blood clotting and in angiogenesis
3	K 00650	AAA52471	Transcription factor
4	M 29039	AAA59198	Transcription factor, Proto-oncoprotein with similarity to Jun that participates in AP-1 transcriptional activation and in malignant transformation, when coexpressed with Jun may act as an inhibitor of Jun-mediated activation and transformation
5	S 59049	AAB26289	GTPase activating protein, Regulator of G-protein signaling 1, one of a family of proteins that negatively regulates G protein-coupled receptor signaling by binding to G-protein alpha subunits and stimulating their intrinsic GTPase activity, expressed specifically in B lymphocytes
6	X 68277	CAA48338	Protein phosphatase, Non-receptor protein-tyrosine phosphatase that has similarity to dual specificity serine/tyrosinephosphatases, upregulated by oxidative stress and heat shock
7	M 36820	AAA63183	Ligand, Chemotactic agent for human polymorphonuclear leukocytes, has strong similarity to human GRO1, which is a mitogenic factor involved in inflammatory processes and that has growth-regulatory properties

Table 3: (continued)

No	mRNA Accession-No	Protein Accession-No	Biochemical Function of the „COPD GENE"
8	M 36821	AAA63184	Ligand, Paralog of human GRO1, which is a mitogenic factor involved in inflammatory processes and that has growth-regulatory properties
9	X02530	CAA26370	Protein kinase, IFN-gamma-inducible protein, a member of the C-X-C chemokine family with homology to platelet factor 4 and beta-thromboglobulin
10	L 20971	AAA03593	Hydrolase, cAMP-specific phosphodiesterase that is sensitive to the antidepressant rolipram
11	L 22214	AAA17544	Receptor (signalling), Adenosine A1 receptor, a G protein-coupled receptor that selectively binds adenosine, stimulates cell death of thymocytes and phagocytosis
12	U 67784	AAB16913	Receptor (signalling), Protein with strong similarity to murine Cmkor1, which is a member of the G protein-coupled receptor family and related to chemokine receptors of the CXC family
13	V 01512	CAA24756	Transcription factor, Transcription factor that forms a heterodimer with c-Jun and stimulates transcription of genes containing activator protein-1 (AP-1) regulatory elements,
14	AL 021154, X66924	CAA48862	Transcription factor; Inhibitor or repressor, Member of the Id helix-loop-helix family of proteins that negatively regulate cell differentiation by inhibiting helix-loop-helix transcription factors via a dominant-negative mechanism
15	M 20681	AAB61083	Major Facilitator Superfamily; Transporter, Facilitated glucose transporter
16	L 06797	AAA03209	Receptor (signalling), CXC chemokine receptor (fusin), a G protein-coupled receptor that binds CXC cytokines and mediates intracellular calcium flux
17	M 60278	AAA35956	Ligand, Heparin-binding EGF-like growth factor that binds to EGF receptors and TM4 superfamily members
18	M 96233	AAA57346	Transferase, Member of the mu class of glutathione S-transferases, a family of detoxification enzymes that catalyzesthe conjugation of glutathione to electrophilic compounds

Table 3: (continued)

No	mRNA Accession-No	Protein Accession-No	Biochemical Function of the „COPD GENE“
19	U03688	AAA19567	Oxidoreductase; Transporter, Cytochrome P450 1B1
20	AF 052124	AAC28619	Extracellular matrix protein
21	M 28225	AAA60308	Ligand, Human JE gene encoding a monocyte secretory protein
22	L 23808	AAA58658	Protease, Matrix metalloprotease that degrades elastin
23	AA 883870	XP_008846	Hydrolase, Apolipoprotein CII, cofactor for lipoprotein lipase which hydrolyzes triglyceride-rich lipoproteins
24	AF 012024	AAB88672	Adhesin/agglutinin, Protein that interacts with the cytoplasmic domain of beta1 integrin and has a role in cell signaling
25	AB 013382	BAA34369	Protein phosphatase, Dual specificity protein phosphatase, a constitutively expressed cytosolic protein that selectively dephosphorylates and inactivates mitogen-activated protein kinase
26	D 26443	BAA05462	Transporter, High affinity sodium-dependent glutamate/aspartate transporter, important for reuptake of glutamate in excitatory neurotransmission
27	X16302	CAA34373	Structural protein, Insulin-like growth factor binding protein that binds to and modulates insulin-like growth factor activity
28	M26683	AAA96708	Ligand, Cytokine A 2, a chemotactic factor for monocytes

Table 4: Primer and Probe Sequences for TaqMan Analyses

Genes, up-regulated in COPD patients

No	*Accession-No*	PRIMER AND PROBES SEQUENCES 5' > 3'	
1	L 13463	Probe forward primer reverse primer	ACGTTCAGACCACCTATTCCCTTCTTGCG ATTCCTGCTTTTCAGAGTGAGACA GAACAGCTTCCCTCACTGTGTACA
2	X 14787	Probe forward primer reverse primer	TCGCCAGGCCCTGTGGTGG AGGGCGTTCCGCAGGT GGGCCTTTCTGTGAAAGTTGTAA
3	K 00650	Probe forward primer reverse primer	TTCCCAGCATCATCCAGGCCCAG CGAGCCCTTTGATGACTTCCT GGCTCCCAGTCTGCTGCATA
4	M 29039	Probe forward primer reverse primer	CGCGTCTCTCAAGCTCGCCTCTTC CAGCTACTTTTCTGGTCAGGGC ACAATCAGGCGTTCCAGCTC
5	S 59049	Probe forward primer reverse primer	ACAGTGGAGCTCTGTATTAGAAAGCCCCTCA GAGGAAGATACTGTGGTACTGTCATAAAA GAGTTACCTGGCCTTCCCAGTT
6	X 68277	Probe forward primer reverse primer	CCAACTTCAGCTTCATGGGCCAGCT GCAGAGGCGAAGCATCATCT CAGCACCTGGGACTCAAACTG
7	M 36820	Probe forward primer reverse primer	TGCCAGTGCTTGCAGACCCTGC CCTGGCCACTGAACTGCG TGGATGTTCTTGAGGTGAATTCC
8	M 36821	Probe forward primer reverse primer	TGAATTCCCTGCAGTGTCTGCAAGCA ACATTCACACTTTGGATGTTCTTGA GAGCGTCCGTGGTCACTGA

Table 4: (continued)

No	*Accession-No*	PRIMER AND PROBES SEQUENCES 5' > 3'	
9	X 02530	Probe forward primer reverse primer	TTCTGACTCTAAGTGGCATTCAAGGAGTACCTCTC CGATTCTGATTTGCTGCCTTATC GCAGGTACAGCGTACGGTTCT
10	L 20971	Probe forward primer reverse primer	TGATGCGGTCTGTCCATTGCCGATA TTGTCTCCCTGCTGGAAAAATT TGAGCAACCCCACCAAGTC
11	L 22214	Probe forward primer reverse primer	CCGCATCCAGAAGTTCCGCGTC GCAACTCGGCCATGAACC GGTCATTCCAAATCTTAAGGAAGG
12	U 67784	Probe forward primer reverse primer	CCGGTCCTTCTACCCCGAGCACA CTGCGTCCAACAATGAGACCTA CCGATCAGCCACTCCTTGAT
13	V 01512	Probe forward primer reverse primer	AGGACCTTATCTGTGCGTGAAACACACCA CCTAGAGGGTTCCTGTAGACCTAGG AGTCCTTGAGGCCCACAGC
14	AL 021154	Probe forward primer reverse primer	CGCCGCCTTGGCATAGTTTGGA TCCAAGGAGACCAGAAGACCA GTGTGGAAGGAGTGGCTGCT
15	M 20681	Probe forward primer reverse primer	TAGATCTGGAAAGGACGGCGTCATGGA GCAGGCACACGGTGCA GCAGGCTCGATGCTGTTCAT
16	L 06797	Probe forward primer reverse primer	CTCCAAAGGAAAGCGAGGTGGACATTCAT GGGTCCAGCCTCAAGATCC ACTTGAAGACTCAGACTCAGTGGAAA
17	M 60278	Probe forward primer reverse primer	CGTGTCCTCTCCCTGCCAAGTCTCAG GTCAAAGTGTAACAGATATCAGTGTCTCC TACAGGCATGGAAGCCCAAC
18	M 96233	Probe forward primer reverse primer	CCTACAATGATGCAGCACTTCTCACAGTTCC GAAGCCAGAATACTTGGAGGAACT ATACGGTGGAGGTCAAGGACAT

Table 5

Genes, down-regulated in COPD patients

No	*Accession-No*	PRIMER AND PROBES SEQUENCES 5' > 3'	
19	U 03688	Probe forward primer reverse primer	ATTCCTCATGCCACCACTGCCAACA TCCAGCTTTGTGCCTGTCAC GGGAATGTGGTAGCCCAAGA
20	AF 052124	Probe forward primer reverse primer	TCACCTCACACATGGAAAGCGAGGAGT CCAGTACCCTGATGCTACAGACG TGGCCTTGTATGCACCATTC
21	M 28225	Probe forward primer reverse primer	ACCCAAACTCCGAAGACTTGAACACCCAC TGGACCACCTGGACAAGCA GCTGCAGATTCTTGGGTTGTG
22	L 23808	Probe forward primer reverse primer	CCTTATGGCCAACCTTGCCATCTGG CTGAGAGACCAAAGACCAGTGTTAAT TTTCTGGCTTCAATTTCATAAGCA
23	AA 883870	Probe forward primer reverse primer	ACAGCCGCCCAGAACCTGTACGAGA TCCAGTTACTGGGAGTCAGCAA TCTACAGCGGGCAGGTATGTC
24	AF 012024	Probe forward primer reverse primer	TTCACGATCCAGCACTGTGGCCA AGCAAGTCTGTGGATTCTAGCCTT CCTGAGCTTTTGGTGGAATCTG
25	AB 13382	Probe forward primer reverse primer	ATGCTTGACTTTACCAATTCTGATGACATCTTTACG GAAAATTGTGCTCTGTTGTAATCCA GCCATAACAAGGTCTTAGTGATAATAGTG
26	D 26443	Probe forward primer reverse primer	TGGACTGGTTCCTGGATCGCCTCC CACTGACGACATCACGCTCAT CCCAGTACGTTGGTGGTGGT
27	X 16302	Probe forward primer reverse primer	AACCTCAAACAGTGCAAGATGTCTCTGAACG CTGTGACAAGCATGGCCTGTA GCACTCCCCACGCTGC
28	M 26683	Probe forward primer reverse primer	CCATGGCAGCCCTTTGGTGCA ACCTGCGTTCCTCCTCTAGCT TGGTCTAGAAGTGCAGCCCATT

Table 6

No	Accession No.	TaqMan at SC	Fold Change	**Primer and Probes sequences 5' > 3'**		Forward Concen-tration (nM)	Reverse Concen-tration (nM)	Probe Concen-tration (nM)
1	L13463	YES	1,86	Forward Sequence Reverse Sequence Probe Sequence	GAACAGCTTCCCTCACTGTGTACA ATTCCTGCTTTTCAGAGTGAGACAC ACGTTCAGACCACCTATTCCCTTCTTGCG	900	900	200
2	X14787	YES	20,45	Forward Sequence Reverse Sequence Probe Sequence	TCTTTGGCTACCAGTCCAGCA AAAGGCCCGAGTATCCCTGA CAAGTCACCCAGTCCTACTGGGACACCAA	900	900	200
3	K00650	YES	8,79	Forward Sequence Reverse Sequence Probe Sequence	TCTAGGACTTCTGCACGGACCT AGGTCCGGACTGGTCGAGAT CAGTGCCAACTTCATTCCCACGGTCAC	900	900	200
4	M29039	YES	2,5	Forward Sequence Reverse Sequence Probe Sequence	CAGCTACTTTTCTGGTCAGGGC ACAATCAGGCGTTCCAGCTC CGCGTCTCTCAAGCTCGCCTCTTC	900	900	200
5	S59049	YES	4,79	Forward Sequence Reverse Sequence Probe Sequence	TGGCTGAAGGGAATTAACAGATAGT TTCATGTTAATCCATTCTGAGTCATTT TCACCCAGGGAGCCATACTGGCACAT	900	900	200
6	X68277	YES	1,52	Forward Sequence Reverse Sequence Probe Sequence	GCAGAGGCGAAGCATCATCT CAGCACCTGGGACTCAAACTG CCAACTTCAGCTTCATGGGCCAGCT	900	900	200
7	M36820	YES	2,95	Forward Sequence Reverse Sequence Probe Sequence	TGCTGCTCCTGCTCCTGG AGGTGAATTCCCTGCAGGGT CACTGAACTGCGCTGCCAGTGCTTG	900	900	200
8	M36821	YES	2,01	Forward Sequence Reverse Sequence Probe Sequence	GAGCGTCCGTGGTCACTGA ACATTCACACTTTGGATGTTCTTGA TGAATTCCCTGCAGTGTCTGCAAGCA	900	900	200
10	L20971	YES	1,39	Forward Sequence Reverse Sequence Probe Sequence	TGCCAGCTATGTGGTAGGGC GTCAACCGTGGGATTTCCTG CCTGGCCTTTCAGTTACTTGAGTTTGGAGTC	50	50	175

Table 6: (continued)

No	Accession No.	TaqMan at SC	Fold Change	**Primer and Probes sequences 5' > 3'**		Forward Concen-tration (nM)	Reverse Concen-tration (nM)	Probe Concen-tration (nM)
12	U67784	YES	11,54	Forward Sequence Reverse Sequence Probe Sequence	CGTGCAAAGTCACACACCTCAT GGTGTTGGTGAAGTAGGTGATGG CTTCCTCACGTGCATGAGCGTGGAC	900	900	200
15	M20681	YES	26,96	Forward Sequence Reverse Sequence Probe Sequence	CCTTTGAAGGGCAGGCACA GCAGGCTCGATGCTGTTCAT TCCATGACGCCGTCCTTTCCAGATCTA	900	900	200
16	L06797	YES	2,24	Forward Sequence Reverse Sequence Probe Sequence	CTGAGAAGCATGACGGACAAGTAC AGGGAAGCGTGATGACAAAGAG CTGCACCTGTCAGTGGCCGACCTC	900	900	200
17	M60278	YES	27,48	Forward Sequence Reverse Sequence Probe Sequence	AGGAACAGGGAACATTGGAGCT TTGGTGAGGTGGGTGGGAT AACTGATTACCTGCCAATTGCTACCGAGAAGGT	900	900	200
18	M96233	YES	7,7	Forward Sequence Reverse Sequence Probe Sequence	AGAGGAGGTCGCAGTTCAGC CAGTACCCCAGTGTCATGGACA TCTGCAGAATCGACACCAACCAGCATC	900	900	200

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 December 2002 (05.12.2002)

PCT

(10) International Publication Number
WO 02/097127 A3

(51) International Patent Classification[7]: C12Q 1/68, G01N 33/48, C07K 14/47, A61K 48/00

(21) International Application Number: PCT/EP02/05835

(22) International Filing Date: 28 May 2002 (28.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0113266.1 31 May 2001 (31.05.2001) GB

(71) Applicant *(for all designated States except US)*: **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **OELLERS, Nadja** [DE/DE]; Steinfeldergasse 9, 50670 Köln (DE). **GEHRMANN, Mathias** [DE/DE]; Rosenstr. 68, 47918 Tönisvorst (DE). **KALLABIS, Harald** [DE/DE]; Wolfskaul 8, 51061 Köln (DE). **HALL, Roderick** [GB/GB]; 10 Acacia Close, Woodham, Surrey KT 15 3SJ (GB). **SCHULZE, Thomas** [DE/DE]; Kaiserwaldstr. 2c, 91083 Baiersdorf (DE). **KROEGEL, Claus** [DE/DE]; Falkenweg 18, 07751 Cospeda (DE).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*

Published:
— *with international search report*

(88) Date of publication of the international search report:
21 August 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/097127 A3

(54) Title: GENES AND PROTEINS FOR PREVENTION, PREDICTION, DIAGNOSIS, PROGNOSIS AND TREATMENT OF CHRONIC LUNG DISEASE

(57) Abstract: Genes that are differentially expressed in lung tissue of COPD patients versus lungs of normal people are disclosed. The genes provide novel methods for the prevention, prediction, diagnosis, prognosis and treatment of chronic lung disease.

INTERNATIONAL SEARCH REPORT

International Application No
T/EP 02/05835

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68 G01N33/48 C07K14/47 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
SEQUENCE SEARCH, EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 00 55174 A (HUMAN GENOME SCIENCES INC ;ROSEN CRAIG A (US); RUBEN STEVEN M (US)) 21 September 2000 (2000-09-21) Seq Id Nos 567 and 1507 page 3, line 3 - line 22 page 354, line 27 - line 29 page 365, line 19 -page 366, line 10 page 380, line 13 - line 17 claims 17-20 ---	1-12
X	US 5 955 314 A (GOLI SURYA K ET AL) 21 September 1999 (1999-09-21) Seq Id No 4 column 7, line 45 -column 8, line 13 abstract --- -/--	1-12

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 January 2003	13. 05. 2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	CEDER O.

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/05835

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 854 191 A (SMITHKLINE BEECHAM CORP) 22 July 1998 (1998-07-22) abstract ---	1-12
A	EP 1 043 406 A (TEIJIN LTD) 11 October 2000 (2000-10-11) abstract -----	1-12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 02/05835

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: -
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [X] Claims Nos.: -
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-12 all partly

Remark on Protest

[] The additional search fees were accompanied by the applicant's protest.

[] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claim 12 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.1

Rule 39.1(iv) PCT - Method for treatment of the human or animal body by therapy

Continuation of Box I.2

Present claim 8 relates to the use of a compound defined by reference to a desirable characteristic or property, namely it being identifyable by any of the methods of claims 5-7.

The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to the compounds identified in parts a) - h) of claim 8, and only as far as the search fees have been paid.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: Claims 1-12 all partly

A method for prevention, prediction, diagnosis and treatment of chronic lung disease utilizing the polynucleotide of Seq Id No 1 or the polypeptide of Seq Id No 29, a method for screening for agents which regulates the polypeptide/polynucleotide or a pharmaceutical composition containing them.

Inventions 2-28: Claims 1-12 all partly

Idem as for invention 1 but for polynucleotide 2-28 and polypeptide 30-56, respectively.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
/EP 02/05835

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0055174 A	21-09-2000	AU 3395900 A	04-10-2000
		AU 3617600 A	04-10-2000
		AU 3617700 A	04-10-2000
		AU 3619400 A	04-10-2000
		AU 3619500 A	04-10-2000
		AU 3869400 A	04-10-2000
		CA 2364567 A1	21-09-2000
		CA 2364590 A1	21-09-2000
		CA 2364629 A1	21-09-2000
		CA 2366130 A1	21-09-2000
		CA 2366174 A1	21-09-2000
		CA 2366195 A1	21-09-2000
		EP 1168917 A2	09-01-2002
		EP 1165588 A1	02-01-2002
		EP 1169469 A1	09-01-2002
		EP 1165589 A1	02-01-2002
		EP 1159420 A1	05-12-2001
		EP 1163358 A1	19-12-2001
		JP 2003512815 T	08-04-2003
		JP 2003512816 T	08-04-2003
		WO 0055173 A1	21-09-2000
		WO 0055350 A1	21-09-2000
		WO 0055351 A1	21-09-2000
		WO 0055180 A2	21-09-2000
		WO 0055174 A1	21-09-2000
		WO 0055320 A1	21-09-2000
		US 2003054421 A1	20-03-2003
		US 2002081659 A1	27-06-2002
		US 2002039764 A1	04-04-2002
		US 2002055627 A1	09-05-2002
		US 2002151681 A1	17-10-2002
		US 2002052308 A1	02-05-2002
		US 2002044941 A1	18-04-2002
US 5955314 A	21-09-1999	AU 5238398 A	29-05-1998
		EP 0958363 A1	24-11-1999
		JP 2001527522 T	25-12-2001
		WO 9820128 A1	14-05-1998
		US 2002034777 A1	21-03-2002
EP 0854191 A	22-07-1998	US 6008017 A	28-12-1999
		EP 0854191 A2	22-07-1998
		JP 11042092 A	16-02-1999
		JP 2002191375 A	09-07-2002
		US 6365715 B1	02-04-2002
EP 1043406 A	11-10-2000	AU 747991 B2	30-05-2002
		AU 1682399 A	05-07-1999
		CA 2315218 A1	24-06-1999
		EP 1043406 A1	11-10-2000
		US 6346385 B1	12-02-2002
		CN 1285002 T	21-02-2001
		WO 9931271 A1	24-06-1999